

SELECTED READINGS IN CONSUMER NEUROSCIENCE & NEUROMARKETING

2nd edition



Compiled by
Thomas Zoëga Ramsøy
2014

ABOUT THIS COMPENDIUM

The original purpose of this compendium has been for the use in my own lectures in consumer neuroscience and neuromarketing at the Copenhagen Business School. However, I also recognise that this volume can also be a potentially valuable resource for both newcomers as well as experienced people within this discipline. Neuromarketing is today very much a conglomerate of divergent solutions; hyped up talks; and a mixture of true science and pop science gone terribly wrong. This collection of papers represent my own take on what the basics should entail

This book is also intended as a supplement to my book “Introduction to Neuromarketing & Consumer Neuroscience”, which you can read more about here: <http://neuronsinc.com/publications/introduction-to-neuromarketing-consumer-neuroscience/> (also see next page).

The selection of texts are not intended to be an exhaustive listing of all relevant articles. I have worked from two basic premises: 1) that the article is available freely on the web; and 2) that the article represents some of the leading thoughts (and scholars) in this field.

If you have suggestions or comments, please send me an email at tzramsoy@gmail.com

DISCLAIMER

All materials in this compendium – texts and images – have either been collected from freely available resources, or written by myself. I claim no ownership or rights over these materials. All materials in this compendium – except my own freely distributed materials – can be collected and compiled by any individual. Please note that there may be restrictions on materials in this compendium for sharing, distributing or selling.

If you find that this compendium contains materials that are not permitted for sharing, please send me an email at tzramsoy@gmail.com and I will adjust accordingly.

Happy reading!

All the best,



WHO MADE THIS?



Thomas Zoëga Ramsøy, b. 1973 in Oslo, Norway

Thomas is considered one of the leading experts on neuromarketing and consumer neuroscience, and he is an innovator by heart. With a background in economics and neuropsychology, he holds a PhD in neurobiology from the University of Copenhagen.

Thomas has published extensively on the application of neuroimaging and neurophysiology to consumer behaviour and decision making. He is the Director of the Center for Decision Neuroscience, where his research team uses an eclectic mix of technologies and the sciences of economics, psychology and neuroscience. Beyond this, Thomas is the CEO of Neurons Inc, where he consults companies around the globe on the use of science and technology in business.

More information about Thomas can be found on the following resources:

Professional pages

DNRG CBS – <http://cbs.dk/DNRG>

DNRG HH – <http://drcmr.dk/research/DecisionNeuroscience>

Neurons Inc – <http://neuronsinc.com>

Social media

Twitter – <https://twitter.com/NeuronsInc>

Neurons Inc – <http://NeuronsInc.com>

BrainEthics – <http://brainethics.org>

Societies

Neuromarketing Science & Business Association – <http://www.neuromarketing-association.com>

Society for Mind Brain Sciences – <http://mbscience.org/>

Publications

ResearchGate – https://www.researchgate.net/profile/Thomas_Z_Ramsoy/

NEW BOOK!

How can we study the unconscious drivers of choice?

This is the authoritative introduction and update to neuromarketing and consumer neuroscience, written by one of the leading figures in this field

Get it here: <http://neuronsinc.com/publications/introduction-to-neuromarketing-consumer-neuroscience/>

Advance praise

"Dr. Ramsøy manages a rare balance – a book that goes deep but ranges wide across the full breadth of this complex but fascinating field. (...) If you want to understand how neuromarketing and consumer neuroscience are changing the landscape of market research today, you need to add this ebook to the top of your reading list right away."

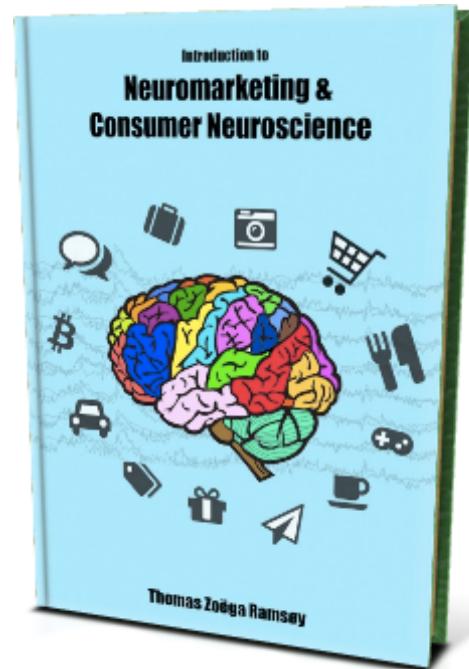
– Stephen Genco, author of Neuromarketing for Dummies

"Thomas' new book is an impressive tour de force! Everything you ever wanted to know about consumer neuroscience together with the very latest on its applications to Neuromarketing. It is hard to believe the book only scratches the surface (Cf. Epilogue). A mini encyclopedia and practitioner's handbook all in one!"

– Franck Sarrazit, Global Director of Brands and Communication, TNS

"This much needed book constitutes a superb outline of neuromarketing's neuroscience foundations as well as some of the recent exciting developments in this field. It is an easy to read account by one of the key contributors to the field."

– Emeritus Professor Richard Silberstein, President of the Neuromarketing Science and Business Association (NMSBA), and Chairman of Neuro-Insight Pty Ltd



About the book

How do we make decisions on what to buy and what to pay for it? Why are we affected by brands and pricing when making our choices or just experiencing something?

Traditional approaches to such questions have relied on the behavioural and social sciences. However, today we see a dramatic shift in our understanding of consumption behaviours. Recent advances in modern neuroscience, and how it combines with economics and psychology, have allowed us to study of how different brain functions serve consumer behaviour. A commercial industry is emerging that offers novel ways to assess consumer attention, emotion and memory.

This book, written by Dr. Thomas Zoëga Ramsøy – one of the leading figures in neuromarketing and consumer neuroscience – offers a comprehensive insight into the workings of the brain and its mind, and how this knowledge can inform our understanding of consumption behaviours.

The book offers both basic and front-end academic insights, and includes chapters on sensation and perception; attention and consciousness; emotion and feeling; memory and learning; motivation and preference; and decision making. It also offers up to date and comprehensive insight about how the tools of neuroscience can be applied to assess consumer cognition and emotion.

This book works as a landmark for this emerging academic and commercial disciplines, and to become a standard book of reference, just as the textbooks by Kotler and Keller have been for advertising and marketing.

Find out more: <http://neuronsinc.com/publications/introduction-to-neuromarketing-consumer-neuroscience/>

CONTENTS

SELECTED READINGS IN CONSUMER NEUROSCIENCE & NEUROMARKETING

Overview & Introduction

1. Phan – Neuromarketing: Who Decides What You Buy?
2. Hannaford – 'Neuromarketing': can science predict what we'll buy?
3. Nobel – Neuromarketing: Tapping Into the 'Pleasure Center' of Consumers
4. Plassmann, Hilke, Thomas Zoëga Ramsøy, and Milica Milosavljevic. "Branding the Brain: A Critical Review and Outlook." *Journal of Consumer Psychology* 22, no. 1 (2012): doi: 10.1016/j.jcps.2011.11.010.
5. Rangel, Antonio, Colin Camerer, and P Read Montague. "A Framework for Studying the Neurobiology of Value-based Decision Making." *Nat Rev Neurosci* 9, no. 7 (2008): doi: 10.1038/nrn2357.
6. Ariely, Dan, and Gregory S Berns. "Neuromarketing: The Hope and Hype of Neuroimaging in Business." *Nat Rev Neurosci* 11, no. 4 (2010): doi:10.1038/nrn2795.
7. Yoon, Carolyn, Richard Gonzalez, Antoine Bechara, Gregory S Berns, Alain A Dagher, Laurette Dubé, Scott A Huettel, and others. "Decision Neuroscience and Consumer Decision Making." *Marketing Letters* 23, no. 2 (2012): doi:10.1007/s11002-012-9188-z.
8. Lee, Nick, Amanda J Broderick, and Laura Chamberlain. "What Is "neuromarketing"? A Discussion and Agenda for Future Research." *Int J Psychophysiol* 63, no. 2 (2007): doi: 10.1016/j.ijpsycho.2006.03.007.

Sensory Neuromarketing

1. Krishna, Aradhna, and Maureen Morrin. "Does Touch Affect Taste? The Perceptual Transfer of Product Container Haptic Cues." *Journal of Consumer Research* 34, no. 6 (2008): 807-818.
2. Morewedge, Carey K, Young Eun Huh, and Joachim Vosgerau. "Thought for Food: Imagined Consumption Reduces Actual Consumption." *Science (New York, N.Y.)* 330, no. 6010 (2010): doi:10.1126/science.1195701.
3. Morrin, Maureen, and S Ratneshwar. "Does It Make Sense to Use Scents to Enhance Brand Memory?" *Journal of Marketing Research* 40, no. 1 (2003): 10-25. <http://www.jstor.org/stable/30038832>.
4. Ramsøy, Thomas Z, Morten Friis-Olivarius, Catrine Jacobsen, Simon B Jensen, and Martin Skov. "Effects of Perceptual Uncertainty on Arousal and Preference Across Different Visual Domains." *Journal of Neuroscience, Psychology, and Economics* 5, no. 4 (2012): doi:10.1037/a0030198.

Attention, Awareness & Consciousness

1. Baars, Bernard J, Thomas Z Ramsøy, and Steven Laureys. "Brain, Conscious Experience and the Observing Self." *Trends in Neurosciences* 26, no. 12 (2003): doi:10.1016/j.tins.2003.09.015.
2. Baars, Bernard J, Stan Franklin, and Thomas Zoëga Ramsøy. "Global Workspace Dynamics: Cortical "binding and Propagation" Enables Conscious Contents." *Front Psychol* 4 (2013): doi: 10.3389/fpsyg.2013.00200.
3. Bargh, John A. "Losing Consciousness: Automatic Influences on Consumer Judgment, Behavior, and Motivation." *Journal of Consumer Research* 29, no. 2 (2002).
4. Chartrand, Tanya L, Joel Huber, Baba Shiv, and Robin J Tanner. "Nonconscious Goals and Consumer Choice." *Journal of Consumer Research* 35, no. 2 (2008): doi:10.1086/588685.
5. Simonson, Itamar. "In Defense of Consciousness: The Role of Conscious and Unconscious Inputs in Consumer Choice." *Journal of Consumer Psychology* 15, no. 3 (2005): 211-217.

6. Milosavljevic, Milica, Vidhya Navalpakkam, Christof Koch, and Antonio Rangel. "Relative Visual Saliency Differences Induce Sizable Bias in Consumer Choice." *Journal of Consumer Psychology* 22, no. 1 (2012): doi:10.1016/j.jcps.2011.10.002.
7. Milosavljevic, M, and M Cerf. "First Attention Then Intention." *International Journal of Advertising* 27, no. 3 (2008): 381-398.
8. Ramsøy, Thomas Zoëga, and Martin Skov. "Brand Preference Affects the Threshold for Perceptual Awareness." *Journal of Consumer Behaviour* 13, no. 1 (2014): doi:10.1002/cb.1451.

Emotions & Feelings

1. Dmochowski, Jacek P, Matthew A Bezdek, Brian P Abelson, John S Johnson, Eric H Schumacher, and Lucas C Parra. "Audience Preferences Are Predicted by Temporal Reliability of Neural Processing." *Nature communications* 5 (2014): doi:10.1038/ncomms5567.
2. Salimpoor, Valorie N, Mitchel Benovoy, Kevin Larcher, Alain Dagher, and Robert J Zatorre. "Anatomically Distinct Dopamine Release During Anticipation and Experience of Peak Emotion to Music." *Nature neuroscience* 14, no. 2 (2011): 257-262.
3. Plassmann, Hilke, John O'Doherty, Baba Shiv, and Antonio Rangel. "Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness." *Proceedings of the National Academy of Sciences of the United States of America* 105, no. 3 (2008): doi:10.1073/pnas.0706929105.
4. Groeppel-Klein, Andrea. "Arousal and Consumer In-store Behavior." *Brain Res Bull* 67, no. 5 (2005): doi:10.1016/j.brainresbull.2005.06.012.
5. McClure, Samuel M, Jian Li, Damon Tomlin, Kim S Cypert, Latané M Montague, and P Read Montague. "Neural Correlates of Behavioral Preference for Culturally Familiar Drinks." *Neuron* 44, no. 2 (2004): doi:10.1016/j.neuron.2004.09.019.
6. Dolcos, Florin, Kevin S LaBar, and Roberto Cabeza. "Dissociable Effects of Arousal and Valence on Prefrontal Activity Indexing Emotional Evaluation and Subsequent Memory: An Event-related FMRI Study." *NeuroImage* 23, no. 1 (2004): doi:10.1016/j.neuroimage.2004.05.015.

Learning & Memory

1. Van Dillen, Lotte F, Dirk J Heslenfeld, and Sander L Koole. "Tuning Down the Emotional Brain: An fMRI Study of the Effects of Cognitive Load on the Processing of Affective Images." *NeuroImage* 45, no. 4 (2009): doi:10.1016/j.neuroimage.2009.01.016.
2. Chessa, A G, and J M J Murre. "A Neurocognitive Model of Advertisement Content and Brand Name Recall." *Marketing Science* 26, no. 1 (2007): doi:10.1287/mksc.1060.0212.
3. Bray, Signe, Antonio Rangel, Shinsuke Shimojo, Bernard Balleine, and John P O'Doherty. "The Neural Mechanisms Underlying the Influence of Pavlovian Cues on Human Decision Making." *The Journal of Neuroscience : the official journal of the Society for Neuroscience* 28, no. 22 (2008): doi:10.1523/JNEUROSCI.0897-08.2008.
4. Tom, Gail, Carolyn Nelson, Tamara Srzentic, and Ryan King. "Mere Exposure and the Endowment Effect on Consumer Decision Making." *The Journal of Psychology* 141, no. 2 (2007): 117-125.

Consumer Preference: Wanting & Liking

1. Santos, José Paulo, Daniela Seixas, Sofia Brandão, and Luiz Moutinho. "Investigating the Role of the Ventromedial Prefrontal Cortex in the Assessment of Brands." *Frontiers in neuroscience* 5 (2011): doi:10.3389/fnins.2011.00077.
2. Berns, Gregory S, and Sara E Moore. "A Neural Predictor of Cultural Popularity." *Journal of Consumer Psychology* 22, no. 1 (2012): doi:10.1016/j.jcps.2011.05.001.

3. Tusche, Anita, Stefan Bode, and John-Dylan Haynes. "Neural Responses to Unattended Products Predict Later Consumer Choices." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30, no. 23 (2010): doi:10.1523/JNEUROSCI.0064-10.2010.
4. Pessiglione, Mathias, Predrag Petrovic, Jean Daunizeau, Stefano Palminteri, Raymond J Dolan, and Chris D Frith. "Subliminal Instrumental Conditioning Demonstrated in the Human Brain." *Neuron* 59, no. 4 (2008): doi:10.1016/j.neuron.2008.07.005.
5. Plassmann, Hilke, John O'Doherty, and Antonio Rangel. "Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 27, no. 37 (2007): doi:10.1523/JNEUROSCI.2131-07.2007.
6. Sescousse, Guillaume, Jérôme Redouté, and Jean-Claude Dreher. "The Architecture of Reward Value Coding in the Human Orbitofrontal Cortex." *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30, no. 39 (2010): doi:10.1523/JNEUROSCI.3501-10.2010.

Decision Making

1. Ravaja, Niklas, Outi Somervuori, and Mikko Salminen. "Predicting Purchase Decision: The Role of Hemispheric Asymmetry Over the Frontal Cortex." *Journal of Neuroscience, Psychology, and Economics* (2012) doi:10.1037/a0029949.
2. De Martino, Benedetto, Dharshan Kumaran, Ben Seymour, and Raymond J Dolan. "Frames, Biases, and Rational Decision-making in the Human Brain." *Science (New York, N.Y.)* 313, no. 5787 (2006): doi:10.1126/science.1128356.
3. Deppe, Michael, Wolfram Schwindt, Anna Pieper, Harald Kugel, Hilke Plassmann, Peter Kenning, Katja Deppe, and E Bernd Ringelstein. "Anterior Cingulate Reflects Susceptibility to Framing During Attractiveness Evaluation." *Neuroreport* 18, no. 11 (2007): doi:10.1097/WNR.0b013e3282202c61.
4. Plassmann, Hilke, Peter Kenning, Michael Deppe, Harald Kugel, and Wolfram Schwindt. "How Choice Ambiguity Modulates Activity in Brain Areas Representing Brand Preference: Evidence From Consumer Neuroscience." *Journal of Consumer Behaviour* 7, no. 4-5 (2008): doi:10.1002/cb.257.
5. Lee, Leonard, On Amir, and Dan Ariely. "In Search of Homo Economicus: Cognitive Noise and the Role of Emotion in Preference Consistency." *Journal of Consumer Research* 36, no. 2 (2009): doi:10.1086/597160.
6. Ramsøy, Thomas Zoëga, and Martin Skov. "How Genes Make Up Your Mind Individual Biological Differences and Value-based Decisions." *Journal of Economic Psychology*

Social Behavior

1. Lee, Victoria K, and Lasana T Harris. "How Social Cognition Can Inform Social Decision Making." *Frontiers in neuroscience* 7 (2013): doi:10.3389/fnins.2013.00259.
2. Schulz, Jonathan F, Urs Fischbacher, Christian Thöni, and Verena Utikal. "Affect and Fairness: Dictator Games Under Cognitive Load." *Journal of Economic Psychology* (2012).
3. Raine, Adrian, and Yaling Yang. "Neural Foundations to Moral Reasoning and Antisocial Behavior." *Soc Cogn Affect Neurosci* 1, no. 3 (2006): doi:10.1093/scan/nsl033.
4. Haselton, Martie G, Mina Mortezaie, Elizabeth G Pillsworth, April Bleske-Rechek, and David A Frederick. "Ovulatory Shifts in Human Female Ornamentation: Near Ovulation, Women Dress to Impress." *Horm Behav* 51, no. 1 (2007): doi:10.1016/j.yhbeh.2006.07.007.
5. Spezio, Michael L, Antonio Rangel, Ramon Michael Alvarez, John P O'Doherty, Kyle Mattes, Alexander Todorov, Hackjin Kim, and Ralph Adolphs. "A Neural Basis for the Effect of Candidate Appearance on Election Outcomes." *Soc Cogn Affect Neurosci* 3, no. 4 (2008): doi:10.1093/scan/nsn040.

6. Kosfeld, Michael, Markus Heinrichs, Paul J Zak, Urs Fischbacher, and Ernst Fehr. "Oxytocin Increases Trust in Humans." *Nature* 435, no. 7042 (2005): 673-676.

Consumer Aberrations

1. Hollander, E, and A Allen. "Is Compulsive Buying a Real Disorder, and Is It Really Compulsive?" *American Journal of Psychiatry* 163, no. 10 (2006): 1670-1672.
2. Ramsøy, Zuraigat, Jacobsen, Bagdziunaite, Skov & Bechara (in progress) Arousal, Executive Control and Decision Making in Compulsive Buying Disorder
3. Black, D W. "A Review of Compulsive Buying Disorder." *World Psychiatry* 6, no. 1 (2007): 14.
4. Black, Donald W, Martha Shaw, and Nancee Blum. "Pathological Gambling and Compulsive Buying: Do They Fall Within An Obsessive-compulsive Spectrum?" *Dialogues Clin Neurosci* 12, no. 2 (2010): 175-85.

Neuroethics of Neuromarketing

1. Murphy, Emily R, Judy Illes, and Peter B Reiner. "Neuroethics of Neuromarketing." *Journal of Consumer Behaviour* 7, no. 4-5 (2008): doi:10.1002/cb.252.
2. Weisberg, Deena Skolnick, Frank C Keil, Joshua Goodstein, Elizabeth Rawson, and Jeremy R Gray. "The Seductive Allure of Neuroscience Explanations." *Journal of cognitive neuroscience* 20, no. 3 (2008): 470-477.
3. Wilson, R, J Gaines, and R P Hill. "Neuromarketing and Consumer Free Will." *Journal of Consumer Affairs* 42, no. 3 (2008): 389-410.

OVERVIEW AND INTRODUCTION



Neuromarketing and Consumer Neuroscience are novel academic and commercial disciplines, but the topic has already been debated for many years. While many articles have focused on the coming neuromarketing trend, only today we see the actual realization of the use of neuroscience tools in consumer insights and marketing.

We should distinguish between *neuromarketing*, which is a commercial use of neuroscience tools to improve consumer insights and marketing effects; and *consumer neuroscience*, which is more an academic exercise, where the aim is to use neuroscience to better understand consumer psychology and behavior.

A central overview paper is Plassmann, Ramsøy & Milosavljevic (2012). This paper both reviews the field as a whole, but also suggests a consumer/branding model that can be used as a map when reading other materials.

Neuromarketing: Who Decides What You Buy?

Vicky Phan

People who have found themselves indulging in clothing trends, jiving to mainstream music, or frequenting the local Starbucks can see that companies spend billions a year researching how to perpetuate such conformity. What people may not know is that the advertising itself is becoming far more scientifically advanced. Neuromarketing is an emerging branch of neuroscience in which researchers use medical technology to determine consumer reactions to particular brands, slogans, and advertisements. By observing brain activity, researchers in lab-coats can predict whether you prefer Pepsi or Coke more accurately than you can. Critics have already begun to denounce the idea for its intrusiveness; however, though the field is already highly controversial, there is no doubt that its continuing development will ultimately have a profound impact on consumerism and the overall study of human behavior.

In America's capitalist society, advertisements drive our everyday lives. While the idea of actual 'mind control' may seem far-fetched and unrealistic, the fact remains that the

Malls are a prime example of where plenty of Neuromarketing takes place, such as here in Plaza Blok M. Mall in Jakarta, Indonesia. Reproduced from [8].



“ Advertising is becoming
more scientifically
advanced ”

marketing industry has had a firm grasp over the American perception of everything from smoking to sex education. Our current concept of marketing, with its image-based ads, department store window displays, and catchy TV jingles, actually did not exist before the mid-1900s. Starting in the 1950s, fast food industries teamed up with processed food companies to shape the concept of what we now understand to be McDonald's and Burger King 'cuisine' [1]. In the 1980s, the invention of cable TV, VCRs, and remote controls revolutionized the advertising world, as it allowed the media to become much more easily accessible to average families [2]. These developments soon allowed advertising executives to cater to the public's general interests and subconscious desires.

Over time, the marketing industry has learned to exploit our responses to a wide variety of images and concepts. It is not difficult, however, to recognize and understand the methodology behind these marketing campaigns. The strategic placement of Victoria's Secret models into Super Bowl halftime commercials has an obvious sexual appeal. Celebrities are paid to endorse particular products, since their personal testimonies make any company just seem better. Even the catchiness of a jingle makes us more likely to pause when we see a bag of Kit Kats or Goldfish crackers. But somehow, despite the almost laughably obvious marketing methods, we still respond positively to popular brands and catchy slogans—tools crafted purposely by marketing executives to catch our attention. This tendency to gravitate toward familiar symbols and phrases is the driving force behind the concept of neuromarketing. Scientists are focusing on these natural inclinations, using brain imaging techniques to gauge consumer reactions and expand upon more common, traditional methods, such as surveys and focus groups [3].

There are multiple types of brain-imaging technologies used in current neuromarketing studies: fMRI (functional magnetic resonance imaging), QEEG (quantitative electroencephalography), and MEG (magnetoencephalography). However, the fMRI method is currently the most popular amongst marketing companies, since it utilizes mainstream technology to produce clear images of real-time brain activity [4]. As an imaging technique, the process also translates results more easily into layman's terms: rather than presenting data in strings of incomprehensible numbers, fMRI technology gives people the opportunity to actually visualize the activity patterns in their brains [5].

fMRI works by gauging amounts of hemoglobin, the oxygen-carrier on red blood cells, in certain parts of the body.



An example of a functional MRI Scanner: UC Berkeley's Varian 4T. Reproduced from [9].

For mental imaging, the machine “measures the amount of oxygenated blood throughout the brain and can pinpoint an area as small as one millimeter” [6]. The harder a specific area of the brain is working, the more oxygen it requires; so when the fMRI machine scans the brain, it picks up on the areas with concentrated amounts of hemoglobin and displays

Smidts in 2002, and the general premise of the research was not widely recognized until the first neuromarketing conference in 2004. However, the potential results and subsequent discoveries about human responses to the media are causing this infant branch of science to rapidly gain popularity [4].

The infamous “Pepsi vs. Coca-Cola” experiment, in which scientists studied the motivation behind brand preferences, was what first put early neuromarketing in the spotlight. The researchers observed that although Pepsi and Coke are essentially identical, people often favor one over the other. They subsequently sought to investigate how cultural messages work to guide our perception of products as simple as everyday beverages [7].

The experiment was simple: there were two taste tests—one blind and one in which subjects knew which beverage was which—and the researchers observed the corresponding brain activity. When volunteers were unaware of which brand they were drinking, the fMRI showed activation in the ventromedial prefrontal cortex, a basic “reward center,” when they drank Pepsi. However, when the subjects knew which soda was which, the scans showed brain activity in the hippocampus, midbrain, and dorsolateral prefrontal cortex (which are centers for memory and emotion), in favor of Coke. So essentially, people actually liked the taste of Pepsi, but they were more inclined to believe that they preferred Coke, based off of nostalgia and emotional connections. From these results, the researchers determined that “a preference for Coke is more influenced by the brand image than by the taste itself” [4].

“ Researchers can predict whether you prefer Pepsi or Coke ”

them as regions of high mental activity on the computer screen. These computer images are what researchers use to identify the parts of the brain being utilized.

For neuromarketing, scientists use fMRI to observe areas of the brain that respond to consumer-based stimuli, such as particular brands, price ranges, and even taste preferences [4]. The researchers have found that the regions in the brain corresponding to the prediction of gain and loss (the nucleus accumbens and the insula, respectively) are indicators of behavior and reaction to finances and economics [3]. In other words, we make our decisions based on cursory judgments of whether we will gain or lose money when purchasing a product.

Though fMRI technology was first used for marketing purposes in the late 1990s, the actual term “neuromarketing” was only just coined by Erasmus University’s Professor Ale

The outcome of these studies is intriguing and even a bit entertaining; however, upon a second glance, it can also be alarming. The fact that a series of ads could actually cause your brain to believe something that contradicts what the rest of your body thinks is unnerving, to say the least. Because of this, there is a growing amount of controversy surrounding the subject of neuromarketing.

“ The amount of research we have today is still minimal ”

One of the more paranoid views on this subject is that people may eventually fall victim to an uncontrollable force compelling them to think or act a certain way. While it is still too early for anyone to make definitive legal restrictions on the technology, people are already anxious about its subliminal undermining of free will. Commercial Alert, an organization protesting the development of neuromarketing, has expressed concern over the use of medical technology for advertising purposes, claiming that brain scans “subjugate the mind and use it for commercial gain” [6]. The group has argued that any power-hungry neuroscientist could use these studies to manipulate the public’s desire for specific products, or that the research could be used in the realm of politics and propaganda, dragging us down a slippery slope toward totalitarianism and war [6].

On the other hand, more optimistic observers contend that the studies could in fact be beneficial for our society. For example, neuromarketing has the potential to be a great boon to public service industries by helping them understand how to improve anti-drug or anti-smoking campaigns [3]. By utilizing these new advancements in neuroscience, we could educate the public more effectively; we would know how to better present information to inattentive children, how to best impact teenagers having unprotected sex, and how to inform the public about conserving energy. The road toward understanding consumer responses opens paths to understanding human behavior in general, which could be invaluable to the development of our global community.

Despite the ongoing debate about the ethics of neuromarketing, the amount of research we have today is still minimal, and the results are leading researchers to believe that nobody currently has the power to fully alter our personal opinions and preferences. Most professionals are presently under the impression that this field is underdeveloped and that researchers are hyping it up using neuroscience, a current ‘hot topic,’ to elicit extra funding [3]. However, though there isn’t much evidence so far to prove that the imaging studies will have a drastic effect on consumers, researchers agree that even a slight edge in the competition

References:

1. Spring, J. Educating the consumer-citizen: a history of the marriage of schools, advertising, and media. Mahwah: Lawrence Erlbaum Associates, Inc.; 2003.
2. Fox, S. The mirror makers: a history of American advertising and its creators Edition 1997. New York: Morrow, 1984.
3. Schnabel, J. Neuromarketers: the new influence-peddlers? The Dana Foundation. 25 Mar 2008 [cited 2009 Oct 26]. Available from: <http://www.dana.org/news/features/detail.aspx?id=11686>.
4. Bridger D, Lewis D. Market researchers make increasing use of brain imaging. ACNR. 2005; 5(3): 36-7.
5. Bloom, P. Seduced by the flickering lights of the brain. Seed Magazine. 2006 Jun 27 [cited 2010 Jan 7]. Available from: <http://seedmagazine.com/content/article/>



Do you really prefer the taste of Coke over Pepsi, or do you just prefer the brand? Reproduced from [10].

to win the public’s attention would be worth the cost for many advertisers.

Like all new scientific advancements, neuromarketing is thus far merely a research tool. Marketing expert Martin Lindstrom views the area of study as “simply an instrument used to help us decode what we as consumers are already thinking about when we’re confronted with a product or a brand” [6]. In either case, the studies would reveal more intimate details about human thought-processing and decision-making on a broader scale.

So the question remains: Is neuromarketing a step forward in understanding the human mind, or is it an invasive marketing ploy geared toward demolishing privacy and personal opinion? As of right now, nobody seems to be sure. Though there is always the possibility that this technology could be exploited for immoral purposes, one could say that any scientific discovery has the same potential for misuse in the wrong hands. The best way to limit the media’s influence is to educate ourselves about the science and to be more deliberate with our decisions; a well-educated consumer is less likely to make rash judgments based on unfounded claims. Still, knowing that companies have people researching how our minds work probably won’t stop most of us from pining after all of the latest products —we will always have commercialism to thank for that. ■

Vicky Phan is an undergraduate from University of California, San Diego.

-
6. Lindstrom, M. Buyology: Truth and Lies about Why We Buy. New York: Doubleday; 2008.
 7. McClure, SM, Li J, Tomlin D, Cybert KS, Montague LM, Montague PR. Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron*. 2004; 44: 379-387.
 8. CC-BY, Jonathan McIntosh. Available from: http://commons.wikimedia.org/wiki/File:Mall_culture_jakarta01.jpg
 9. PD. Available from: <http://commons.wikimedia.org/wiki/File:Varian4T.jpg>
 10. CC-BY, Ralph Unden. Available from: <http://www.flickr.com/photos/ralphunden/245218460/>

'Neuromarketing': can science predict what we'll buy?

Advertisers have long used science to peer into consumers' brains; today 'neuromarketing' has given them the power to delve into our subconscious, finds Alex Hannaford



Neuromarketing Illustration Photo: NEIL WEBB

By Alex Hannaford

7:00AM BST 13 Apr 2013

In 2008, a team of scientists in Germany published a study showing how the brain unconsciously prepares our decisions: that several seconds before we consciously decide what we're going to do, its outcome can be predicted by looking at unconscious activity in our grey matter.

The researchers, from the Max Planck Institute in Leipzig, told participants in the study that they could freely decide if they wanted to press a button with their left or right hands,

whenever they wanted, but they had to remember at which time they felt they had made up their mind. They found that it was possible to predict from their brain signals which option they would choose seven seconds before they consciously made their decision.

"It's all very Minority Report," Steve Sands says, referring to the Tom Cruise film in which a special police department known as "PreCrime" tracks down criminals based on knowledge provided by psychics. "But we're not too far from that now."

In fact, it's incredible how close Sands is. For the past 20 years, from his lab in El Paso, Texas, he's been using technology to look inside our heads and show what consumers really feel, as opposed to what marketers think we feel. Using EEG tests (essentially a plastic swimming cap complete with electrodes to measure brain signals), functional magnetic resonance imaging (fMRI, which measures brain activity by looking at changes in blood flow), and eye-tracking technology, neuromarketing, as it's known, has completely revolutionised the worlds of advertising and marketing.

Sands sits opposite me, looking relaxed in a white shirt and jeans. In his office there's a framed 1995 cover of Newsweek on the wall with the headline: "The new science of the brain: why men and women think differently".



"That's the first cap I made," he says of the EEG-outfitted woman on the cover. Sands used to work with rhesus monkeys in the psychology department at the University of Texas. When his lab closed down he started Neuroscan, which became one of the world's largest suppliers of EEG equipment to research scientists. After selling Neuroscan, Sands and his team started to use the same machines to look at the brain's response to advertising.

He recently finished a one-and-a-half year project for POPAI, an international trade association, for which Sands's researchers used eye-tracking and EEG technology to gain insight into shopper habits. The results were fascinating. Forget scrawled shopping lists on the back of an envelope: Sands found that the vast majority (76 per cent) of US grocery shoppers make their purchase decisions in-store, and that shoppers using non-cash payment methods are most likely to make impulse purchases. So shelf-placement and in-store marketing are more

crucial than ever.

Sands's team would pop a pair of eye-tracking glasses on their volunteers (which were in turn wired up to a MacBook Air, carried in a rucksack), then send them off around the store to do their shopping. The researchers then waded through three terabytes of data and analysed 80,000 eye movements from the shoppers that agreed to take part in the study. Sands says a single eye movement takes just 200 milliseconds, the time a product in store gets to persuade a shopper to buy it. "And it only takes one eye movement to change their behaviour," he says.

The researchers noted what Sands calls "approach-avoidance" taking place in the sweets and chocolate aisle, and that the eyes sought out the shopper's favourite sweets, even though they may have decided not to succumb to buying them. "Twenty per cent of eye movements relate to what you're going to buy. The rest are alternatives," Sands says. "We'd watch them pick up a packet of doughnuts, put them back, then walk away. Some came back later and put them in their shopping cart."



Sands says one interesting observation was that while the fizzy drinks aisle was the most organised in the entire supermarket, the sweets and chocolate aisle was the least: what Sands describes as "a potpourri of different sizes, shapes and brands that makes a lot of noise".

"Our brain is looking for something simple, and it's happiest when it finds what it's looking for," he says. "Candy is unusually noisy. The industry doesn't organise itself as well as the canned drinks one does. Visual clutter really does matter. All you're doing is frustrating the brain."

The neuromarketing industry isn't just interested in what makes shoppers choose the

products they do in the supermarket. Much of their work is done before they've even walked through the door.

Each year, Sands Research screens the commercials that have aired during the Super Bowl, the FA Cup of American football, to a test group of around 30 people. As in other tests, his team wires each person up to an EEG machine to monitor their brain signals, and each wears a pair of eye-tracking glasses so the Sands researchers can see what, specifically, they're focusing on.

Super Bowl ads are the most sought-after and expensive slots in the industry. In 2011, among the companies vying for hearts, minds, and cold, hard cash, were Coca-Cola and Volkswagen, both of which came out with brilliant spots. Coca-Cola's featured two border guards in different military uniforms at some godforsaken desert outpost, who bond over a bottle of Coke.

As for Volkswagen, their ad for the VW Passat saw a pint-size Darth Vader walking down the hallway of his suburban home, attempting to use "The Force" on his parents' exercise bike, the washing machine — even the family dog. When his father arrived home in his Passat, the boy was almost ready to admit defeat: he ran outside and tried one last time to use his powers on the car, while inside the house his dad saw what he was trying to do and started the car's ignition with the remote control. The boy turned around, astonished that The Force worked.



Of all the ads Sands has ever tested, The Force was, to use the American vernacular, off the charts, achieving the highest "neuro-engagement score" ever. Adweek named it 2011's best commercial; it won two Gold Lions at Cannes. Before the game even began it had attracted 12 million YouTube views. At the time of writing it's had almost 58 million.

The man behind the advert was Deutsch LA's Doug Van Praet. He says Sands's research

demonstrated that The Force ad had an inordinate capacity to engage the brain. "It galvanised our attention, our engagement and our emotion, and it turned out to be a very powerful predictor of end-market performance." As he writes in his book *Unconscious Branding: How Neuroscience Can Empower (and Inspire) Marketing*, "it drove significant increases in purchase consideration, upped traffic to the VW website by half, and contributed to a hugely successful sales year for the brand."

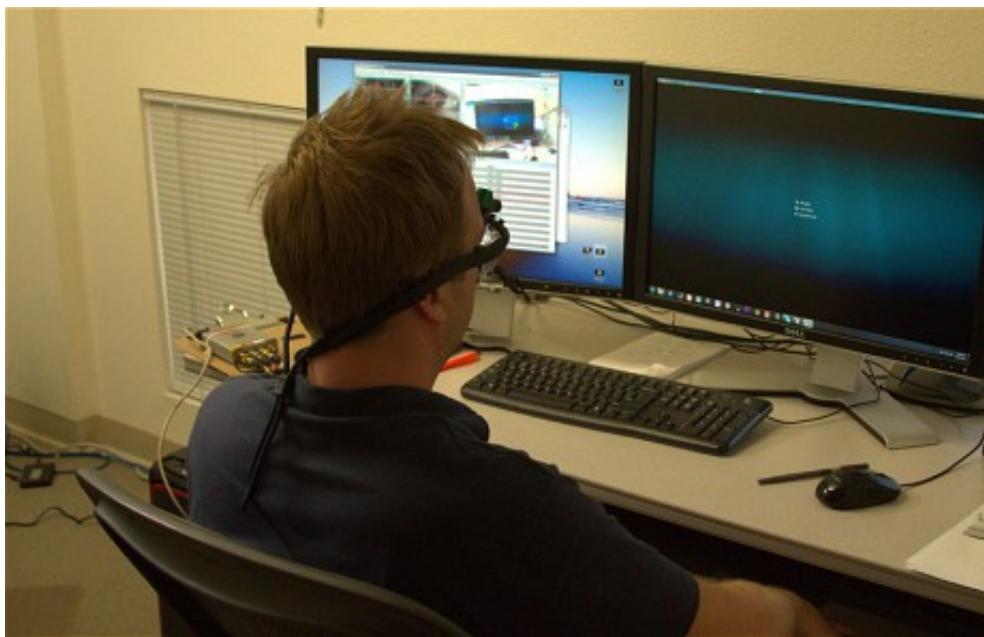
Amazingly, it's taken big business 20-plus years to realise how effective neuromarketing can be. British psychologist Dr David Lewis-Hodgson founded one of the earliest firms, Mindlab International, in the UK in 1988. One of his early "products" was Mindscan, a piece of software which measured the brain's responses to marketing messages, working on the premise: "What can't speak, can't lie."

Hollywood's interest was piqued by a 2012 study by Innerscope Research. They showed 40 film trailers to more than 1,000 people, measuring their heart rate, breathing, how much they sweat and motion responses – as well as what they focused on using eye-tracking technology.

Using the results, they found they could predict box office hits. According to Fast Company magazine, "If a film's trailer fails to reach a specific emotional engagement threshold (65), it will very likely generate less than \$10 million in revenue on opening weekend." But a film whose trailer exceeds an engagement threshold of 80 "will very likely earn more than \$20 million the first weekend". Studios such as Fox and Paramount have now started taking neuromarketing very seriously.

And in January, the research agency Millward Brown announced its clients Unilever and Coca-Cola would be using facial coding technology – where emotions are tracked in facial expressions – in all their advertising testing in 2013. According to the company, this would "automatically interpret viewers' emotional and cognitive states, moment by moment".

As well as gauging an audience's reaction to Super Bowl commercials and tracking shoppers in supermarkets, Sands Research also tests ads before they've aired, letting agencies know what works and what doesn't. "By looking at the EEG readout we can tell whether they're disengaged or engaged," Sands says. "And we've found that storyline wins every time. If you want to lose someone's attention, have several storylines in your commercial."



In a small office adjoining his, Sands sits me in front of two computer monitors and hooks me up to a pair of eye-tracking glasses. On the left-hand screen I can see my eyes, with a target indicating where my pupils are as they flit from left to right; on the right-hand screen is an ad for the Hyundai Sonata hybrid. The places on the screen where my eyes land are denoted by a frenetic green dot that jumps around at incredible speeds. When we play back the recording of my viewing session, I seem to have focused on exactly what the advertisers intended: initially the various characters in the commercial, but by far the longest spell is devoted to eyeing up the car itself.

Only when he places an EEG cap on the head of his test subjects, however, can Sands really tell whether they like what they're seeing. He says he once looked at an ad for a telecoms company and the panel he showed it to had such a negative reaction he had to tell the agency to think again. "There were two competing storylines," he says. "Young creatives think they're multitaskers, and this influences how they design things. And it doesn't work. Simplicity wins every time."

Using EEG, Sands records the electrical activity of the brain along the scalp. In ads that really engage an audience, a large portion of the cerebral cortex, the part of the brain that plays a role in memory, attention, awareness and thought, is activated. Sands says that during portions of an ad that "work", the frontal lobe, which deals with emotion and processes information, lights up. On the computer screen, Sands sees a line, much the same as you see on a heart monitor, which shows the exact moments during the commercial that different parts of the brain are engaged.

Sands's company has also been using EEG technology to gauge taste and smell. "From our experience, people usually tell market researchers what they think they want to hear. We're

social animals and we don't want to offend anyone." But by looking at the brain's response to scents, Sands can tell exactly which fragrance you prefer.

Van Praet acknowledges that neuromarketing is not without its pitfalls; that in studying the human brain, we have to be comfortable with paradox and contradictions. For example, he says you can like an ad and it can create a positive emotion, but if it doesn't leave you with an appropriate and corresponding set of associations and emotions for that product, it's no use to the company trying to sell it. He gives an example: Quiznos, the US sandwich chain that now has some locations in the UK. "They ran an ad that featured cartoonlike rodents, and it was funny as heck," Van Praet says. "It was very likeable and engaging and people remembered it. But there's a bad association between rats and food. It wasn't very successful."

Follow SEVEN on Twitter: @TelegraphSeven

© Copyright of Telegraph Media Group Limited 2013

Forbes



HBS Working Knowledge, Contributor
Leading research and ideas from Harvard Business School faculty.

LEADERSHIP | 2/01/2013 @ 4:48PM | 9,312 views

Neuromarketing: Tapping Into the 'Pleasure Center' of Consumers

It's easy for businesses to keep track of what we buy, but harder to figure out why. Enter a nascent field called neuromarketing, which uses the tools of neuroscience to determine why we prefer some products over others. Harvard Business School marketing professor Uma R. Karmarkar explains how raw brain data is helping researchers unlock the mysteries of consumer choice in this article by Carmen Nobel, which first appeared on the [HBS Working Knowledge website](#).

In the early 1950s, two scientists at McGill University inadvertently discovered an area of the rodent brain dubbed "the pleasure center," located deep in the nucleus accumbens. When a group of lab rats had the opportunity to stimulate their own pleasure

centers via a lever-activated electrical current, they pressed the lever over and over again, hundreds of times per hour, forgoing food or sleep, until many of them dropped dead from exhaustion. Further research found pleasure centers exist in human brains, too.

Most humans are a little more complicated than rats, of course. But we are largely motivated by what makes us feel good, especially when it comes to our purchasing decisions. To that end, many major corporations have begun to take special interest in how understanding the human brain can help them better understand consumers. Enter a nascent but fast-growing field called neuromarketing, which uses brain-tracking tools to determine why we prefer some products over others.

"People are fairly good at expressing what they want, what they like, or even how much they will pay for an item," says Uma R. Karmarkar, an assistant professor at Harvard Business School who sports PhDs in both marketing and neuroscience. "But they aren't very good at accessing where that value comes from, or how and when it is influenced by factors like store displays or brands. [Neuroscience] can help us understand those hidden elements of the decision



(Photo credit: Wikipedia)

process."

To be sure, there is a clear difference between the goals of academia and the goals of a corporation in utilizing neuroscience. For Karmarkar, her work falls into the category of decision neuroscience, which is the study of what our brains do as we make choices. She harbors no motive other than to understand that process and its implications for behavior, and draws on concepts and techniques from neuroscience to inform her research in marketing.

For corporations, on the other hand, the science is a means to an end goal of selling more stuff. But the tools, once restricted to biomedical research, are largely the same. And Karmarkar expects brain data to play a key role in future research on consumer choice.

(In [a recent HBS industry background note on neuromarketing](#), she discusses the techniques that have helped researchers decode secrets such as why people love artificially colored snack food and how to predict whether a pop song will be a hit or a flop.)

Tricks of the trade

When tracking brain functions, neuroscientists generally use either electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) technology. EEG measures fluctuations in the electrical activity directly below the scalp, which occurs as a result of neural activity. By attaching electrodes to subjects' heads and evaluating the electrical patterns of their brain waves, researchers can track the intensity of visceral responses such as anger, lust, disgust, and excitement.

Karmarkar cites the example of junk-food giant [Frito-Lay](#), which in 2008 hired a neuromarketing firm to look into how consumers respond to Cheetos, the top-selling brand of cheese puffs in the United States. Using EEG technology on a group of willing subjects, the firm determined that consumers respond strongly to the fact that eating Cheetos turns their fingers orange with residual cheese dust. In her background note, Karmarkar cites [an article in the August 2011 issue of Fast Company](#), which describes how the EEG patterns indicated "a sense of giddy subversion that consumers enjoy over the messiness of the product."

That data in hand, Frito-Lay moved ahead with an ad campaign called "The Orange Underground," featuring a series of 30-second TV spots in which the Cheetos mascot, Chester Cheetah, encourages consumers to commit subversive acts with Cheetos. (In one commercial, an airline passenger quietly sticks Cheetos up the nostrils of a snoring seatmate. Problem solved.) The campaign garnered Frito-Lay a 2009 Grand Ogilvy Award from the Advertising Research Foundation.

EEG vs. fMRI

Karmarkar notes that EEG and fMRI have different strengths and weaknesses, and that EEG has some limitations in its reach. "The cap of electrodes sits on the surface of your head, so you're never going to get to the deep areas of the brain with EEG," Karmarkar explains.

The fMRI uses a giant magnet, often 3 Teslas strong, to track the blood flow throughout the brain as test subjects respond to visual, audio, or even taste cues. The technology has its own logistical limitations. Running an fMRI scanner costs researchers up to \$1,000 per hour, and studies often use 20-30 subjects, Karmarkar says. And while EEG lets subjects move around during testing, fMRI requires them to lie very still inside a machine that can be intimidating.

"This is a sophisticated piece of medical equipment that exerts a very strong magnetic field at all times, and it's important to be very careful around it," Karmarkar says. "For example, you cannot take metal into a magnet room!"

But fMRI is invaluable to neuroscience and neuromarketing in that it gives researchers a view into the aforementioned pleasure center. "The more desirable something is, the more significant the changes in blood flow in that part of the brain," Karmarkar says. "Studies have shown activity in that brain area can predict the future popularity of a product or experience."

In her note, Karmarkar discusses [research](#) by Emory University's Gregory Berns and Sara Moore, who connected the dots between neural activity and success in the music industry. In a seminal lab experiment, teenagers listened to a series of new, relatively unknown songs while lying inside an fMRI machine. The researchers found that the activity within the adolescents' pleasure centers correlated with whether a song achieved eventual commercial success. The OneRepublic song [Apologize](#) performed especially well in both the brain scans and the market.

"Importantly, Berns and Moore also asked their original study participants how much they liked the songs they heard, but those responses were not able to predict sales," Karmarker's note states, illustrating the marketing value of subconscious cerebral data.

Neuromarketing can provide important but complex data to companies that target a global audience. While product testing may provide similar neural responses in American and Asian subjects, for instance, the marketing implications may be very different.

"Expressions of happiness in some Eastern cultures are expressed as a sense of calm or peace, whereas in some Western cultures, happiness means jumping around with joy and excitement," Karmarkar explains. "So you might get two totally different fMRI results that actually mean the same thing—or you may have two totally different stimuli create the desired effect of profound happiness, but for different reasons. If you get an excited effect in an Eastern market, it may not be a good outcome, even though that was the effect you wanted in a Western market. On the other hand, a sense of peace might be misconstrued as a failure."

Valid concerns

For businesses looking to enlist the services of a neuromarketing company, she advises watching out for consulting firms that claim to offer such services but don't really have the technology or expertise to back up the claim. Rather, look for a company whose employees have a healthy, skeptical respect for neuroscience.

"The rubric for picking a good [firm] is making sure it was started by a

scientist, or has a good science advisory board," Karmarkar says. "This is a field where scientists are very, very skeptical, and we should be. It's easy to feel like you've discovered some big, important truth when you see that the brain has done something that correlates with behavior. And it's just as easy to overstate our conclusions."

For consumers, the idea of giving advertisers additional insight into the subconscious mind might prompt privacy concerns. But Karmarkar says that the research is more about understanding brain waves, not controlling them.

"It's similar to the concerns about genetics," she explains. "People wonder, now that we can map the genome, are we going to manipulate the genome? I think it's a valid and important question to ask. But I don't think it's the direction that companies should take or that academics are taking."

She adds, though, that we need to keep in mind that advertisers have been successfully controlling our brains, to some extent, since long before the existence of EEG or fMRI technology.

"Imagine Angelina Jolie biting into an apple," she says. "It's the juiciest apple ever. She's licking her lips. There's juice running down her chin. Now if I spend some time setting up that scenario and then follow up by asking you to tell me how much you like Mac computers, I *promise* you that you'll rate them more highly than you would have if I hadn't just talked about how great that apple was for Angelina Jolie. So, yes, I just used your brain to manipulate you. Sex sells, and it has since the dawn of time. It sells because it engages that pleasurable reward center of your brain. As academics, neuroscience just helps us to understand how."



About the author: Carmen Nobel is senior editor of *Harvard Business School Working Knowledge*.

This article is available online at:

<http://www.forbes.com/sites/hbsworkingknowledge/2013/02/01/neuromarketing-tapping-into-the-pleasure-center-of-consumers/>

Branding the brain: A critical review and outlook

Hilke Plassmann ^{a,*}, Thomas Zoëga Ramsøy ^{b,c}, Milica Milosavljevic ^d

^a INSEAD, Boulevard de Constance, 77305 Fontainebleau, France, and Decision Neuroscience Group, Cognitive Neuroscience Unit, INSERM, Ecole Normale Supérieure, Paris, France

^b Decision Neuroscience Research Group, Department of Marketing, Copenhagen Business School, Frederiksberg, Denmark

^c Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark

^d Division of Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA

Received 2 February 2011; received in revised form 28 November 2011; accepted 30 November 2011

Available online 21 January 2012

Abstract

The application of neuroscience to marketing, and in particular to the consumer psychology of brands, has gained popularity over the past decade in the academic and the corporate world. In this paper, we provide an overview of the current and previous research in this area and explain why researchers and practitioners alike are excited about applying neuroscience to the consumer psychology of brands. We identify critical issues of past research and discuss how to address these issues in future research. We conclude with our vision of the future potential of research at the intersection of neuroscience and consumer psychology.

© 2011 Society for Consumer Psychology. Published by Elsevier Inc. All rights reserved.

Keywords: Consumer neuroscience; Neuromarketing; Branding; Attention; Memory; Value

Introduction

The application of neuroscience to consumer psychology, and in particular to branding, has gained popularity over the past decade in academic research and business practice: in the last decade the number of publications in top marketing journals and Google references around this topic has grown exponentially and the same holds for the number of neuromarketing companies founded (see Fig. 1).

The birth of the field of consumer neuroscience has generated wide-ranging, ongoing debates of whether this hybrid field benefits its parent disciplines (consumer psychology and neuroscience) and, within them, what forms these benefits might take (Ariely & Berns, 2010; Kenning & Plassmann, 2008; Lee, Broderick, & Chamberlain, 2007; Plassmann, Ambler, Braeutigam, & Kenning, 2007). The goal of consumer neuroscience is to adapt methods and

theories from neuroscience—combined with behavioral theories, models, and tested experimental designs from consumer psychology and related disciplines such as behavioral decision sciences—to develop a neuropsychologically sound theory to understand consumer behavior.

To appreciate the value of combining neuroscience with consumer psychology, it is important to understand the broad range of insights available from neuroscience. Neuroscience is the study of the nervous system that seeks to understand the biological basis of behavior. This range of insights is too broad for the study of consumer psychology, which is why in the following paragraphs we briefly clarify which areas within neuroscience are the most relevant for consumer neuroscience.

Neuroscience research ranges from studying single cells (cellular neuroscience) to studying how different brain areas or complex brain systems, such as the visual system, interact (systems neuroscience). Because of the complexity of consumer behavior,

* Corresponding author.

E-mail addresses: hilke.plassmann@insead.edu (H. Plassmann), tzr.marktg@cbs.dk (T.Z. Ramsøy), mmilosav@hss.caltech.edu (M. Milosavljevic).

URL: <http://www.decisionneuroscience.net> (H. Plassmann).

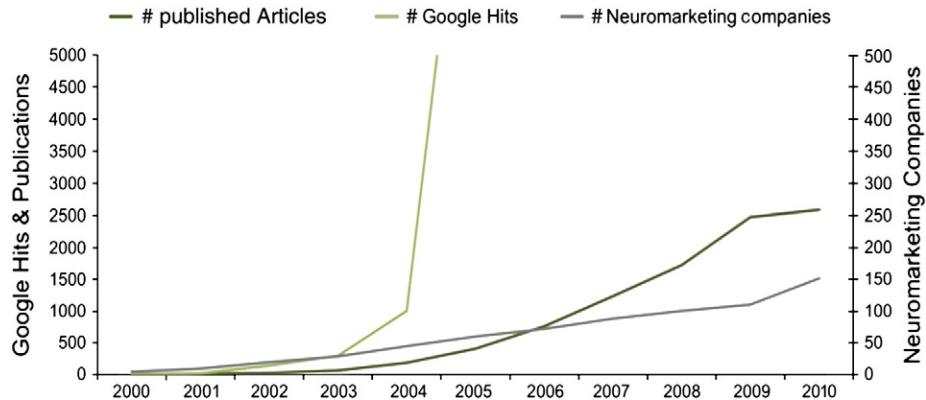


Fig. 1. Growth of research applying neuroscience to marketing over time.

insights from systems neuroscience are crucial for consumer neuroscience, whereas those from cellular neuroscience currently are limited.

Neuroscientists study species ranging from the primitive (such as sea snails, fruit flies, and leeches) to the complex (such as mammals and primates). Most consumer neuroscience studies investigate mental processes in human subjects, but a few selected studies also use non-human primates or small animals such as monkeys as subject populations.¹

Another important distinction is between clinical and non-clinical research in neuroscience. Clinical research, known as neurology, studies how nervous system disorders, trauma, tumors and injuries affect cognition, emotion, and behavior in patients as compared to healthy subject populations. In general, consumer neuroscience studies consumer responses in healthy subject populations.²

A last critical distinction is between consumer neuroscience, which refers to academic research at the intersection of neuroscience and consumer psychology, and neuromarketing, which refers to practitioner and commercial interest in neurophysiological tools, such as eye tracking, skin conductance, electroencephalography (EEG), and functional magnetic resonance imaging (fMRI), to conduct company-specific market research. Neuromarketing has received considerable attention in the corporate world, and the growth of neuromarketing companies over the last decade has been impressive (see Fig. 1).

The goal of this paper is to shed light on what neuroscience can bring to the table to advance our understanding of the consumer psychology of brands. In particular, we aim to provide an overview of the current state of research in this area, identify

critical issues of past research and discuss how to address these issues in future research. We conclude with our vision of the future potential of research at the intersection of neuroscience and consumer psychology.

What is currently done: toward an interdisciplinary understanding of consumer decision making

In this section, we review previous work in neuroscience pertinent to understanding underlying processes involved with brand decisions. We structure the review using a simple consumer decision-making framework based on prior work in consumer psychology (Fig. 2; Kahneman & Snell, 1992; Kahneman, Wakker, & Sarin, 1997; Rangel, Camerer, & Montague, 2008; Wirtz, Kruger, Scollon, & Diener, 2003). We also use this framework to integrate previous consumer neuroscience studies that are directly related to branding questions and to point the way for future applications in consumer research.

The framework divides the stages that are required for brand preference formation over time into four basic components: (1) representation and attention, (2) predicted value, (3) experienced value, and (4) remembered value and learning. Below we explain these basic components and review previous findings on the underlying neuropsychological processes of each of those components. The main brain areas involved with each component of the model are shown in Fig. 3.

Representation and attention

The amount of information consumers are exposed to is enormous, yet our processing capacity is limited. Each second we are exposed to an estimated 11 million bits of information that reach us through all our senses, yet humans are capable of processing only around 50 bits of that information, letting most of the input go by unnoticed (Wilson, 2002). How consumers represent, attend to, and perceive incoming information may have a profound influence on their behavior. In the current section, we discuss representation (i.e., brand identification) and attention.

¹ There are at least two major reasons to study non-human subjects in consumer neuroscience. First, studying animals allows consumer neuroscientists to make causal links between brain areas and specific behaviors. Animal work allows the application of more invasive methods to brain systems that animals and humans have in common. Second, if consumer neuroscience researchers are using evolutionary theories to explain phenomena in consumer behavior such as behavioral biases, using an animal model allows evolutionary inferences (i.e., going back in the evolutionary chain).

² However, there are several reasons to use patient populations in consumer neuroscience. The most prominent one is to use patients with brain lesions to establish causal relationship between brain regions and consumption behavior. At the end of this paper, we will discuss some of these aspects as potential future developments.

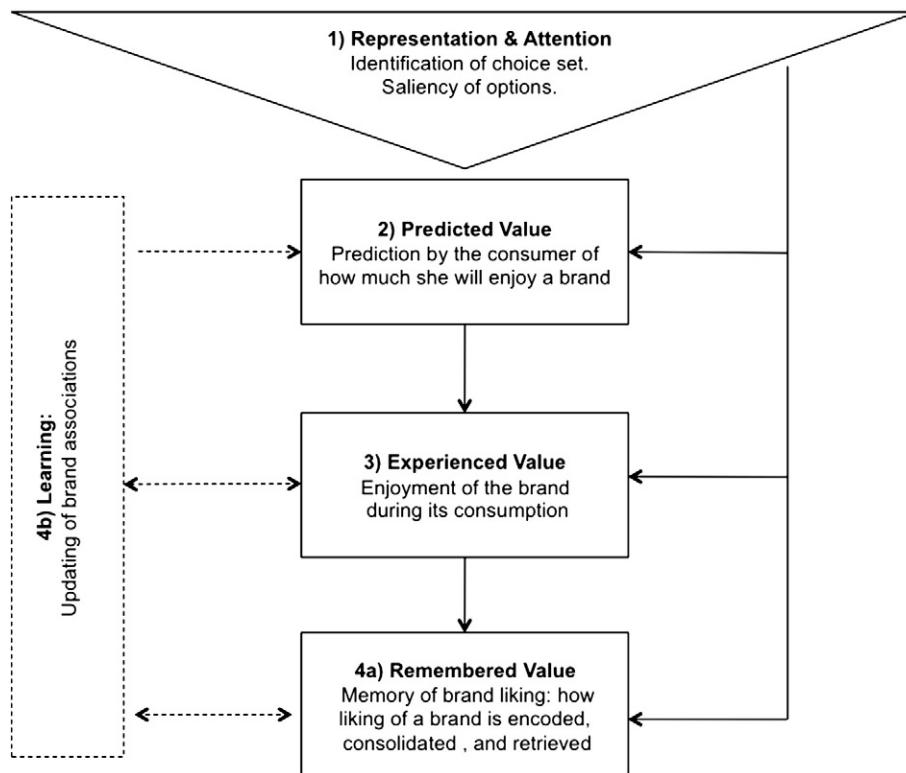


Fig. 2. Value signals important for brand decisions.

Representation The first process in brand decisions involves forming the representation of the choice alternatives—that is, brand identification. This entails processing the incoming information, so that different options for choice are identified (e.g., different beer brands). At the same time, the consumer needs to integrate information on internal states (e.g., thirst level) and external states (e.g., location, social context) that drive attention. For example, when faced with a choice between drinking Heineken or Beck's beer (an incoming information) a consumer's choice is likely to depend on her own level of thirst (an internal state) and what her friend chooses to drink (an external state).

Humans are predominately visual creatures, and most of the incoming information we receive is visual (Koch, 2004). Our visual system contains two cortical routes that are involved with visual processing (see Fig. 3). The dorsal visual pathway is involved with the spatial deployment of attention (the “where/how” pathway) and proceeds from the primary visual cortex V1 in the occipital lobe, through the posterior parietal cortex, to the dorsolateral prefrontal cortex (dIPFC). The ventral visual pathway is responsible for object recognition (the “what” pathway) and originates in V1, then continues to the inferotemporal cortex, and to the ventrolateral PFC.

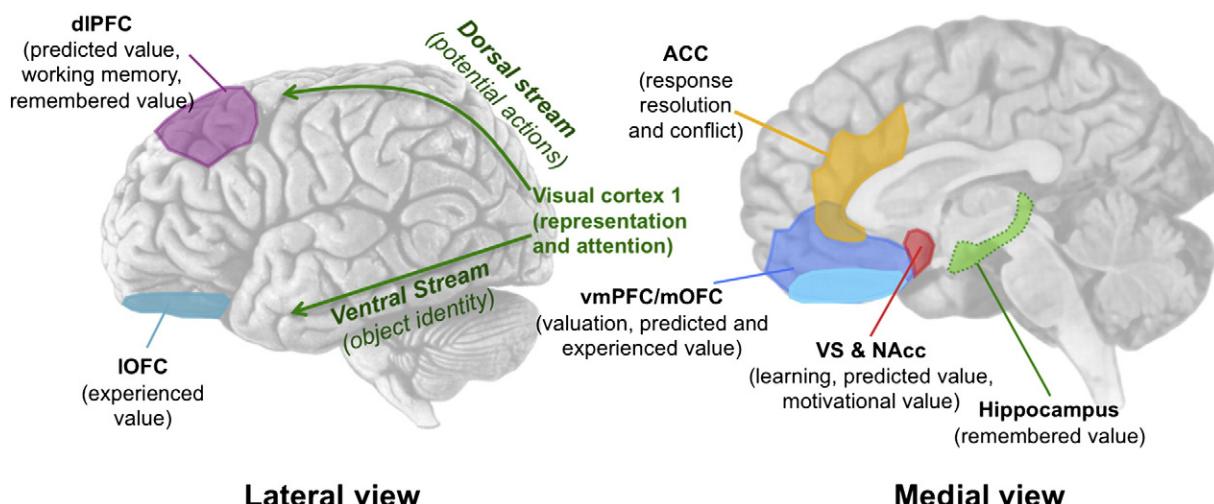


Fig. 3. Overview of prominent brain areas involved in brand decisions. Abbreviations used: ACC = anterior cingulate cortex; dIPFC = dorsolateral prefrontal cortex; IOFC = lateral orbitofrontal cortex; mOFC = medial orbitofrontal cortex; NAcc = nucleus accumbens; vmPFC = ventromedial prefrontal cortex; VS = ventral striatum.

The visual system allows for rapid brand and product identification. A recent magnetoencephalography (MEG) study showed that female participants viewing shoes (compared to motorcycles) had stronger activity in occipitotemporal regions between 130 and 180 ms after image presentation (Junghofer et al., 2010). Similarly, Milosavljevic, Koch, and Rangel (2011) showed that consumers can identify two different food brands and make up their mind about which one they prefer in as little as 313 ms. Furthermore, processes involved in the representation stage need not even be conscious, as recent studies have demonstrated that unconscious processes also shape how we represent our decision-making situations (Chartrand, Huber, Shiv, & Tanner, 2008). One of the key questions at this stage, discussed next, is what consumers pay attention to (i.e., focus on) once they are exposed to a number of rapidly identified choice alternatives (i.e., brands).

Attention Attention is the mechanism responsible for selecting the information that gains preferential status above other available information. Recent review of attention in neuroscience indicates that four conceptual components are fundamental to attention: bottom-up or saliency filters, top-down control, competitive visual selection, and working memory (Knudsen, 2007). We will focus on the first three components and discuss their relevance for research on branding.

Bottom-up or saliency filters automatically select the most important information from all available information. This selection is based on the low-level features of the visual input: colors, luminance, orientation, size, shape, movement, etc. (Itti, Koch, & Niebur, 1998; Wolfe & Horowitz, 2004). Such bottom-up factors have a strong effect on the initial eye movements when consumers are exposed to marketing information: the first four eye-movements are made within the initial 2.5 s of exposure (Leven, 1991). Some higher-level factors are also capable of gaining automatic, preferential access to attention. These include faces, text, novelty, and one's own name.

All of these features are combined in the brain, and preattentional scan paths are created, making a saliency map of the regions in the visual field that are most important and thus most likely to be further processed.

Thus, at the outset of early attention, the decision maker is biased toward salient stimuli (van Zoest, Donk, & Theeuwes, 2004). The salient stimuli will attract the initial eye movements of consumers, and thus may have a profound effect on related consumer behavior.

For example, Pieters and Wedel (2007) showed that ensuring that consumers pay attention to the brand displayed in a print ad is the most effective way to ensure that they will transfer their attention to other elements of the print ad. Further, Milosavljevic and colleagues showed that salient features (i.e., the brightness of the food packaging) influence real food choices (Milosavljevic, Navalpakkam, Koch, & Rangel, 2011). Namely, at fast decision speeds a significant number of food choices were biased toward the food items with brighter

packaging, even when subjects preferred the taste of alternative food options.

There are other automatic biases known to influence what people pay attention to (Glaholt, Wu, & Reingold, 2010). For example, people tend to look toward the upper visual field (Durgin, Doyle, & Egan, 2008) and the right visual field (Efron & Yund, 1996), which may be of importance in the consumer behavior context (e.g., at the point of purchase). In a famous experiment, when five identical stockings were displayed horizontally, subjects were biased toward choosing stockings on the outmost right (Nisbett & Wilson, 1977). Chandon and colleagues showed that only the top-shelf positions carry through to brand evaluation (Chandon, Hutchinson, Bradlow, & Young, 2009). Clearly, products can be placed in locations that are known to attract more attention and will thus be more likely to be chosen by a buyer (Pieters & Warlop, 1999).

Strong location effects were also found when consumers browse websites (Dreze & Hussherr, 2003). The influence of bottom-up factors may be especially strong online, as consumers engage in fast web surfing and often spend very little time on any given page. Systematically manipulating low-level visual features to "guide" viewers' eyes to a webpage's regions of interest is possible by utilizing insights from visual neuroscience. Milosavljevic (2009) used a computer simulation of visual attention to optimize banner ads, and the rest of a website, to make certain brands/banner ads visually salient. This manipulation resulted in an increased liking for the target banner ad, perhaps due to mere exposure effects (Milosavljevic & Cerf, 2008). Recently, a strong bias of looking toward the center of the viewing area (e.g., the center of the computer screen) has been reported (Tatler, 2007). Reutskaya and colleagues showed that an item in the center of the screen was almost 60% more likely to be chosen by a decision maker than similar items displayed at other locations (Reutskaya, Nagel, Camerer, & Rangel, 2011).

Top-down control depends on internal and external states, goals, and expectations. Hence, looking for a can of Coke will enhance processing of red areas in visual input by increasing the neuronal sensitivity for that particular color (Theeuwes, 2010; Treisman & Gelade, 1980; Van der Lans, Pieters, & Wedel, 2008). Expectation can modulate what consumers pay attention to via brain structures that include the dorsolateral cortex (Egidi, Nusbaum, & Cacioppo, 2008). The information that is relevant for goal attainment will be attended to more than irrelevant information. For example, when we are thirsty, we pay more attention to drinks than to other items (Aarts, Dijksterhuis, & De Vries, 2001; Dijksterhuis & Aarts, 2010).

Goals also exert a strong influence on eye-movements and can result in different eye-movement patterns when subjects are exposed to the same visual input (Glaholt et al., 2010; Pieters & Wedel, 2007; Yarbus, 1967). Rosbergen, Pieters, and Wedel (1997) identified tendencies in how individuals scan marketing materials, such as print ads or store shelves. Their work was based on a well-established idea of visual scan paths, that is, the patterns of saccades and fixations across some visual input (Norton & Stark, 1971). They found three types of eye movements that are characteristic of people

examining the ads: scanning (eyes move to headline and pictorial), initial (eyes move to headline, pictorial, and brand), and sustained (eyes move to headline, pictorial, brand, and text). As one might expect, the time spent viewing the ad, the level of involvement, brand attitude, and recall all improved from the first to the third type of viewing. Further, Pieters and Wedel (2007) showed that the informativeness of ads is contingent on the goals consumers pursue while viewing them. For example, in comparison with free viewing of the same ads, consumers spend more time on the text when asked to evaluate the brand, and less time on pictorial elements when asked to learn about the brand.

Visual selection occurs when the most important information from all the areas that are identified as potentially important in preattentive scans (based on the bottom-up input) is chosen. This means that attention is given to a particular location in space. It is believed that as the number of choice options increases, the decision maker becomes more selective in what information he or she encodes, that is, which locations in the scene he or she processes (Payne, Bettman, & Johnson, 1993).

Glaholt et al. (2010) showed that when asked to choose the most expensive of six items (6-alternative-forced-choice, or 6-AFC), subjects were more selective in the processing of stimulus information (i.e., they achieved greater differentiation between individual stimuli via more fixations, longer duration of total fixations, etc.) than when they were asked to choose which of the two sets of three items (2-AFC) was more expensive. Thus, gaze selectivity increases as the number of alternatives increases (Glaholt et al., 2010). Reutskaya et al. (2011) showed that time pressure induced people to shorten the duration of their fixations and to search somewhat longer so as to increase the number of options that are considered before making a choice.

Visual selection and eye movement enhance the quality of incoming information. Gaze bias shows that people spend longer time examining (i.e., fixating on) options that they eventually choose (Glaholt & Reingold, 2009; Krajbich, Armel, & Rangel, 2010; Pieters & Warlop, 1999; Shimojo, Simion, Shimojo, & Scheier, 2003). For example, consumers spent 54% more time looking at the ads of businesses (in a phone directory) that they ended up choosing (Lohse, 1997). It is especially interesting to note that externally manipulating what people look at—for example, by displaying choice options one at a time while manipulating the exposure duration—biases the resulting choices toward the options subjects are exposed to longer (Armel, Beaumel, & Rangel, 2008).

Further, eye movements may be useful in evaluating the effectiveness of brand extensions. Stewart, Pickering, and Sturt (2004) showed that consumers spend 200 ms longer examining implausible brand extensions (they cause immediate disruption of visual processing) compared to plausible brand extensions. The authors propose eye-tracking as a useful tool for determining the extent to which consumers find different brand extensions plausible.

In sum, representation and attention are complex processes that influence all subsequent steps in our brand decisions framework. Theoretical and methodological insights from neuroscience can prove especially useful in allowing consumer researchers to better understand attention and its effects on

branding-related behavior. However, research in this area has received little attention in consumer neuroscience, which offers a lot of potential for future research.

Predicted value

The predicted value of each brand that is available for choice (e.g., Heineken vs. Beck's) represents the consumer's belief about the experienced value of that brand at some time in the future. In other words, the predicted value involves the consumer's evaluation of how much enjoyment she will derive from consuming a Heineken or a Beck's beer.

Previous studies suggest that at least three brain structures might be of particular importance when consumers evaluate predicted values: the striatum, the ventral medial prefrontal cortex (vmPFC), and the dorsolateral prefrontal cortex (dlPFC; see Fig. 3). In the next sections, we first review these previous studies and then review studies that have investigated how branding influences predicted value signals in each respective brain region. For the latter we use Keller's customer-based brand equity framework to categorize the different studies (Keller, 1993). Applying Keller's framework, we distinguish between studies investigating how favorableness, type, and uniqueness of brand associations alter the neural signatures of predicted value (see Table 1). Fig. 4 visualizes the results of the studies listed in Table 1 and shows which brain areas are involved in representing Keller's framework in the brain.

Predicted value signals in the striatum Several studies have used functional magnetic resonance imaging to investigate the predicted value of products or other types of desirable objects such as money. Pioneering work by Knutson and colleagues showed that a structure within the ventral striatum (VS), the nucleus accumbens (NAcc), is involved in encoding anticipated rewards of monetary payoffs (Ballard & Knutson, 2009; Knutson, Adams, Fong, & Hommer, 2001; Knutson & Cooper, 2005; Talmi, Dayan, Kiebel, Frith, & Dolan, 2009) and branded products (Knutson, Rick, Wimmer, Prelec, & Loewenstein, 2007; Knutson et al., 2008).

Two studies investigated how favorableness of brand associations affects predicted value signals in the striatum. In the first one, Schaefer and Rotte (2007a) found that imagining a pleasant experience, such as driving a car of a brand that is linked to favorable brand associations, correlates with activity changes in that brain area. However, it remains unclear what exactly consumers were imagining and whether activity in the striatum is based on the difference in pleasantness of the predicted experience per se or the difference in brand information. This weakness of the study is further confounded by the fact that the more attractive car brands are also more expensive, and driving an expensive car might be a pleasurable experience by itself.

One problem with using a given brain activation (the striatum) to infer a mental process (a pleasurable experience) is the proposed one-to-one relationship between the brain activity and the mental process of interest. Such a "reversed inference" is problematic because one brain area is usually involved in more than one mental process (for a detailed discussion of the "reverse

Table 1
Overview of consumer neuroscience studies directly related to branding.

Branding area	Author	Main question	Method	Main results
Favorability of brand associations	Deppe, Schwindt, Kugel, Plassmann, and Kenning (2005)	What brain areas correlate with brand preference?	fMRI	When choosing between one's favorite brand as compared to a second or lower ranked brand increased activity in the vmPFC and reduced with activity in dlPFC/IFG and visual cortex (cuneus/precuneus) are triggered.
	Deppe, Schwindt, Kramer, et al. (2005)	What are the underlying neural processes of how brand information bias semantic product judgments?	fMRI	Activity in the ACC predicted whether a person is biased by the brand name of a newspaper while evaluating the credibility of a headline.
	Deppe et al. (2007)	What are the underlying neural processes of how brand information bias visual product judgments?	fMRI	Activity in the ACC predicted whether a person is biased by the brand name of a newspaper while evaluating the attractiveness of print advertisements.
	Schaefer and Rotte (2007a)	What are the neural correlates of brand preferences during (imagined) consumption?	fMRI	Imagining driving a car from one's favorite brand correlates with activity changes in the ventral striatum. Activity in this area also correlates with perceived luxury and sportiness of the car.
	Koenigs and Tranel (2008)	What is the role of the vmPFC for how brand information biases preference study judgments?	Lesion study	Patients with damage in the vmPFC were not biased by brand information during blind vs. open tasting of Coke and Pepsi.
	Plassmann et al. (2008)	Does uncertainty modulate the neural signatures of brand preference?	fMRI	Interaction of brand preference with uncertainty of the decision amplifies the neural correlate of brand preference in the vmPFC.
	Erk et al. (2002)	What are the neural correlates of fmri preferences for product types that are vs. low in social status signaling?	fMRI	Sports cars vs. limousines induced increased activity changes in the brain areas involved in reward processing (striatum, vmPFC/mOFC and ACC).
	Schaefer and Rotte (2007b)	Does high social status signaled by brands trigger the same responses than low social status signals?	fMRI	<ul style="list-style-type: none"> • Car brands signaling high social status activated regions in the MPFC and precuneus. • Car brands signaling low social status activated the left superior frontal gyrus and ACC
	McClure et al. (2004)	What are the underlying brain processes of how brand information alters brand evaluations during consumption?	fMRI	<p>Stated preferences for Coke vs. Pepsi did not correlate with revealed preferences in blind tastings</p> <p>Revealed preference correlated with activity changes in the vmPFC/mOFC</p> <p>Knowing you drink Coke vs. not knowing what you drink correlated with activity changes in memory/association areas (hippocampus, dlPFC/SFG). No such difference could be found for the case of Pepsi</p>
	Yoon et al. (2006)	Do brand judgments recruit the same neural networks as judgments about people?	fMRI	Brain areas involved in making judgments about human traits for people do not overlap with brain areas involved in making judgments about human traits for brands.
Brand recall and memories	Schaefer et al. (2006)	What are the neural correlates of brand familiarity?	fMRI	Activity changes in the MFG correlate with familiar vs. unfamiliar brands
	Klucharev et al. (2008)	How does the expertise of an endorser affect brand memory and attitude?	fMRI	Increased brand recall for expert endorsement was related to stronger activation during encoding of memory structures of the left hemisphere, the dlPFC and medial temporal lobe structures, and accompanied by stronger engagement of the bilateral striatum.
	Esch et al. (2012)	What are the neural correlates of brand familiarity and brand "strength"	fMRI	<p>Unfamiliar brand logos vs. "strong" brands induce activity changes in the IFG</p> <p>"strong" vs. unfamiliar brands induce activity changes in the hippocampus and lingual gyrus</p> <p>"strong" vs. "weak" brands induce activity changes in the dlPFC/MFG</p>
Brand loyalty	Plassman, Kenning, and Ahlert (2007)	Do loyal customers recruit other brain areas than disloyal customers during brand choice?	fMRI	Activity in the striatum correlates with brand loyalty to retail brands.

Note: The table includes studies that uncover brain areas involved in different topics related to branding that allow making inferences about locations in the brain. We did not include studies investigating temporal dynamics, such as techniques with a high temporal resolution such as EEG/MEG.

Abbreviations used: ACC = anterior cingulate cortex; dlPFC = dorsolateral prefrontal cortex; IFG = inferior frontal gyrus; mOFC = medial orbitofrontal cortex; vmPFC = ventromedial prefrontal cortex.

inference problem" in consumer neuroscience studies see below). A potentially interesting direction for further studies in this area is to manipulate the expected pleasantness of the consumption experience (e.g., a road trip vs. commuting in heavy traffic) and

investigate how this is altered by brand information while controlling for price levels.

The second study, by Plassmann, Kenning, and Ahlert (2007), found that when choosing between buying identical

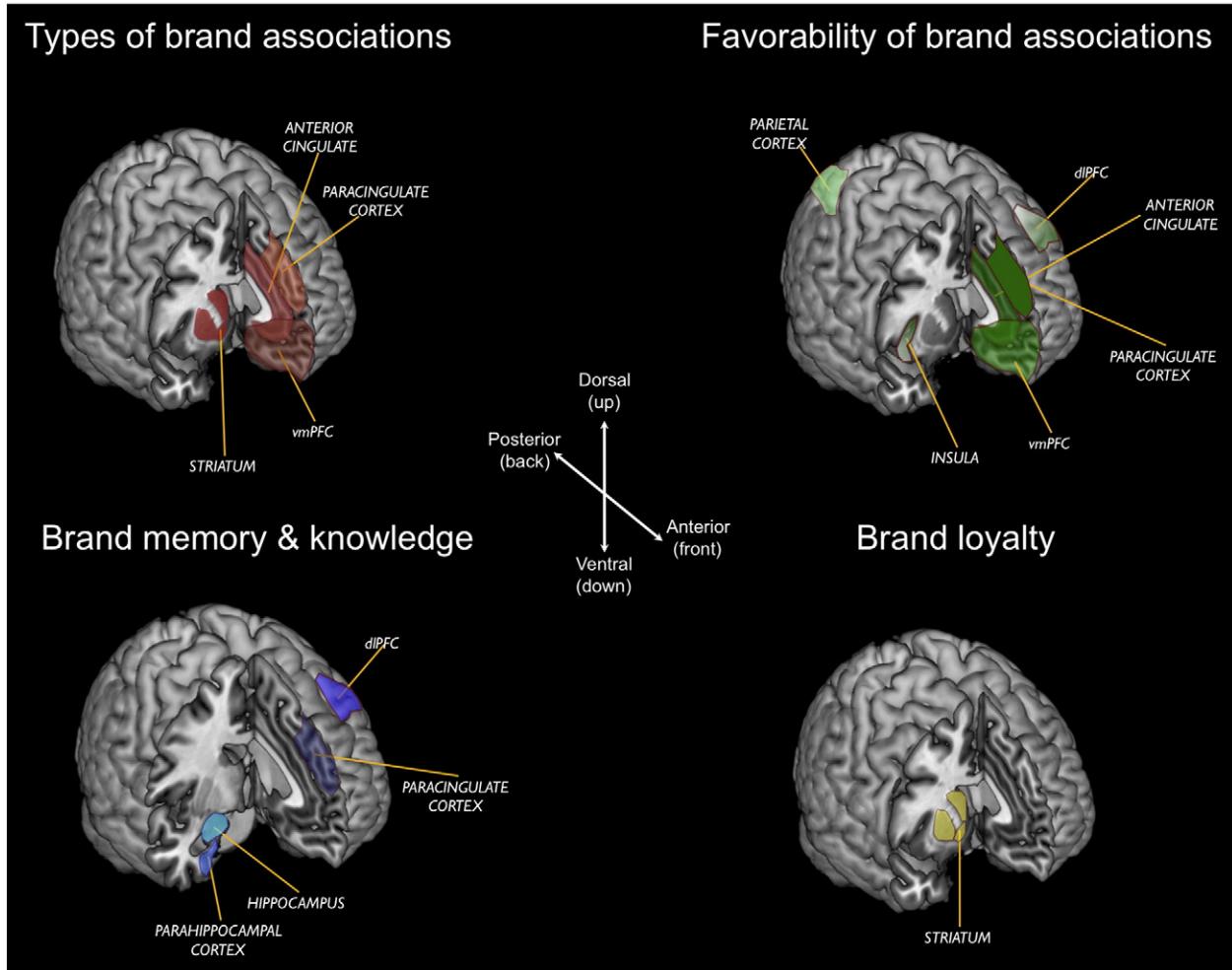


Fig. 4. Overview brain systems involved in the psychology of brands. Note: this figure shows a three-dimensional view of the brain and it is important to note that the front (anterior) to back (posterior) slice view shown differs between the different parts of the figure, i.e. a view showing the hippocampus is more posterior than views showing the striatum; Abbreviations used: dlPFC = dorsolateral prefrontal cortex; vmPFC = ventromedial prefrontal cortex.

clothes at different retail stores (e.g., H&M vs. Zara), customers who are loyal to a store as measured by real purchasing behavior (i.e., amount spent, frequency and recency of purchases based on loyalty card data) show more activation in the striatum compared to customers who are less loyal. Although this study provides an interesting and methodologically valuable link between real-life purchasing behavior outside the lab (i.e., based on scanner data at the point of sale) and brain activation by inviting loyalty card holders to the brain imaging lab, one potential confound of this study is the passive choice paradigm applied in the study. In the passive choice paradigm test persons were not required to respond, i.e. make choices inside the scanner. Instead, behavioral measures were taken outside the scanner. The lack of these response measures results in missing important manipulation checks. This is further linked to the above-mentioned problematic reliance on reverse inference.

To the best of our knowledge, no study to date has investigated the impact of the different types of brand associations or uniqueness of brand associations on predicted value signals in the striatum. This calls for further research in this area.

Predicted value signals in the ventromedial prefrontal cortex (vmPFC) and the dorsolateral prefrontal cortex (dlPFC) Another series of human fMRI studies has studied predicted values using real choices and has found that neural activity in the vmPFC correlates with behavioral measures of consumers' positive and negative predicted values for a range of different branded products (Chib, Rangel, Shimojo, & O'Doherty, 2009; Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Plassmann, O'Doherty, & Rangel, 2007; Plassmann, O'Doherty, & Rangel, 2010) and also in a social context when making decisions about charitable donations (Hare, Camerer, Knoepfle, & Rangel, 2010).

Most of the abovementioned studies found that a brain system consisting of the vmPFC and the dlPFC encodes behavioral preferences (Camus et al., 2009; Plassmann et al., 2007; Plassmann et al., 2010; Wallis & Miller, 2003). Importantly, in Camus et al.'s (2009) study, test subjects were "virtually lesioned" in the dlPFC using transcranial magnetic stimulation (TMS) and showed a change in behavioral measures of predicted values.

Several consumer neuroscience studies have investigated how brand associations alter predicted value signals in the vmPFC and dlPFC as well as in related brain areas. Some of these studies investigated how brand favorability influences neural signatures of predicted value signals; other studies compared how different types of brand associations alter those signals.

A series of studies by Deppe and colleagues investigated how favorability of brand associations influences predicted value signals in the brain. In the first study, the authors found increased neural activity in the vmPFC when the choice set contained the consumer's favorite brand compared to choice sets containing two less preferred brands. They also found that the part of the dlPFC involved in working memory and the part of the visual system involved in object recognition were less active when the choice set contained the consumer's favorite brand compared to a set containing two non-preferred brands (Deppe, Schwindt, Kugel, Plassmann, & Kenning, 2005). A potential confound of this first study is the passive choice design that did not allow the recording of choices and reaction times. In other words, no actual behavioral choices were recorded, but preference rankings were sampled at the end of the experiment outside the scanner. These measures could have served as important manipulation checks and would have avoided having the results rely on reversed inference, discussed below.

In two follow-up studies the authors applied an active choice task and replicated their finding that the vmPFC correlated with favorability of brand associations (Deppe et al., 2005; Deppe et al., 2007). Interestingly, in these follow-up studies, the authors showed that the degree to which a brain area involved in selective attention and conflict monitoring (the anterior cingulate cortex, ACC; see Fig. 3) is correlated with the degree of how much consumers' judgments are biased by brand associations (Deppe, Schwindt, Kramer, et al., 2005; Deppe et al., 2007). In other words, these two studies suggest that ACC activity predicts individual differences of how much brand associations influence consumers' judgments.

A more recent study by Esch et al. (2012) also investigated how favorability of brand associations influences brain activity during brand decisions. They found that the part of the dlPFC involved in predicted value encoding is more active when consumers are exposed to "strong" vs. "weak" brands. They also found that exposure to "weak" vs. "strong" brands leads to more activity in the insula, the brain area previously found to encode disgusting, painful, or more generally intense and arousing emotional experiences. However, because the reported results are based on a very low statistical threshold not corrected for multiple comparisons (i.e., .005) or for cluster levels (i.e., small volume corrections or region of interest analysis) that are typically not reported as core findings in neuroscience journals, the study by Esch et al. awaits further empirical validation.

Several studies have investigated how different types of brand associations influence predicted value signals in the vmPFC, dlPFC, and related areas. Most of the studies look at brand associations linked to cultural influences and social status. Studies by Erk, Spitzer, Wunderlich, Galley, and Walter

(2002) and Schaefer and Rotte (2007a) found that exposure to branded products associated with high social status induces activity changes in the vmPFC, ACC, PFC and striatum. For both studies it remains unclear whether the type of brand association (i.e., high vs. low social status) or how much people like the brand (i.e., favorability of brand association) is driving the results because the experimental design does not allow these two factors to be dissociated. As a result, both studies also rely on reverse inference.

A study by Yoon, Gutchess, Feinberg, and Polk (2006) investigated brand personality associations. The authors compared whether judgments about personality attributes of people are represented in the same neural system as judgments about personality attributes of brands and whether this differs when these judgments refer to the self or others. They found that brain areas involved in making judgments about human traits for people do not overlap with brain areas involved in making judgments about human traits for brands. These first findings challenge the view that we associate brands with personalities and are able to form relationships with brands the same way we form relationships with people (Aaker, 1997; Aaker & Fournier, 1995; Aggarwal, 2004; Fournier, 1997; Swaminathan, Page, & Gurhan-Canli, 2007) and call for further research.

To the best of our knowledge, no study has looked at the impact of the uniqueness of brand associations on predicted value signals in the vmPFC and dlPFC to date. This calls for further research in this area.

Experienced value

Experienced value is based on the pleasure derived from consuming a brand. According to early notions of utility or value, experienced value is the "true value" that should matter the most for value-based decision making (Kahneman et al., 1997). Experienced value consists of the (a) valence and (b) intensity of the consumption experience. In this section, we first review general and branding-related neuroscientific research investigating valence and intensity of experienced values and then review the neural basis of a concept that connects brain systems involved in representing predicted and experienced value, namely motivational value.

Valence The neural bases of computations made by the evaluation system during the consumption experience are beginning to be understood. Human fMRI studies have shown that activity in the orbitofrontal cortex (OFC), in particular its medial parts (see Fig. 3), at the time a reward is being enjoyed correlates with subjective reports about the pleasantness or valence of the experience. This has been shown for olfactory experiences (Anderson et al., 2003; Kringlebach, O'Doherty, Rolls, & Andrews, 2003; McClure et al., 2004; Small, Zatorre, Dagher, Evans, & Jones-Gotman, 2001; Small et al., 2003), musical rewards (Blood & Zatorre, 2001), visual rewards (Aharon et al., 2001; Kirk, Skov, Hulme, Christensen, & Zeki, 2009), pleasantness of touch (McCabe, Rolls, Bilderbeck, & McGlone, 2008), and even secondary rewards such as money (Breiter,

Aharon, Kahneman, Dale, & Shizgal, 2001; Knutson, Fong, Adams, Varner, & Hommer, 2001; Knutson, Fong, Bennett, Adams, & Hommer, 2003). Moreover, the activity in the OFC is reduced when consumers are fed to satiety on a specific food (O'Doherty et al., 2000).

Taken together, these findings suggest that the medial OFC might be an area where positive experienced values are computed. Other studies have found that brain areas that receive inputs from the OFC areas, such as the ventral striatum and the pregenual cingulate cortex (Grabenhorst, Rolls, & Bilderbeck, 2008; McCabe et al., 2008; Rolls, Grabenhorst, & Franco, 2009; Rolls & McCabe, 2007), are also correlated with sensory pleasantness.

An interesting open question is which neural systems encode negative experiences. Several studies have found that unpleasantness of taste might be correlated with brain activity in the lateral OFC and left dorsal anterior insula/operculum (Small et al., 2001; Small et al., 2003). O'Doherty and colleagues found that the size of abstract punishments (i.e., losing money) activated lateral parts of the OFC (O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). One problem in investigating negative experience is to dissociate it from intensity. This problem arises due to the negativity bias of intensity: negative experiences are usually also perceived to be more intense and thus are often confounded (Small et al., 2003), in particular for visual stimuli such as facial or object attractiveness.

Using a different methodological approach to investigate positive vs. negative emotional experiences, neuromarketing studies are based on the idea that there is a left-right asymmetry of the frontal electroencephalography (EEG) signals (Davidson, Ekman, Saron, Senulis, & Friesen, 1990). These and related studies suggest that relatively greater activity in the left frontal region is associated with either positive emotional experience or the motivational drive to approach an object (Harmon-Jones, 2003). Although there are strong correlations between frontal EEG asymmetry and personality traits, the degree to which the asymmetry changes from one moment to another is questionable. Some studies have applied this approach to measure moment-to-moment fluctuations in emotional responses to advertisements without accounting for autocorrelations in time or multiple statistical comparisons (Ohme, Reykowska, Wiener, & Choromanska, 2009). However, the validity of such approaches is unclear, as hemispheric asymmetry is also an index of working memory load (Habib, Nyberg, & Tulving, 2003; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994). Further research to investigate the neural representation of positive vs. negative experienced values is needed.

Several recent human fMRI experiments have provided novel insights into how marketing actions such as branding might alter the properties of the experienced value signals. For example, one study showed that activity in the medial OFC in response to an odor depended on whether subjects believed that they smelled cheddar cheese or body odor (de Araujo, Rolls, Velazco, Margot, & Cayeux, 2005). In another study, activity in the medial OFC in response to the

consumption of wine depended on quality beliefs about its price (Plassmann, O'Doherty, Shiv, & Rangel, 2008). Yet another study found that experienced values of works of art, and accompanying engagement of the medial OFC, depended on whether the subjects believed they were created by an expert (i.e., an artist) or by a non-expert (i.e., the experimenter; Kirk et al., 2009). Together, these findings suggest that the experienced valuation system is modulated by higher cognitive processes that determine expectancies and beliefs—a phenomenon recently referred to as the “placebo effects of marketing” actions (Shiv, Carmon, & Ariely, 2005; Waber, Shiv, Carmon, & Ariely, 2008) or “expectation bias” (Plassmann & Niessing, 2010).

To date there is only one study that has investigated how favorable brand associations alter experienced value signals. McClure et al. (2004) investigated differences in brain activity during consumption of sodas when the subjects knew they were drinking Coke or Pepsi vs. when they did not know which brand they were consuming. Unbeknownst to the subjects, they were consuming Coke and Pepsi in both conditions (brand-cued and non-brand-cued trials). The study showed that the experienced value signals depended on brand associations. In particular, the authors found that subjects' knowing they were drinking Coke vs. not knowing what they were drinking correlated with activity changes in their memory/association areas (hippocampus, dlPFC/SFG). No such difference could be found for Pepsi.

An interesting follow-up of this experiment would be to have four different types of trials. In two types of trials Coke would be administered, once cued with a Coke logo and once cued with a Pepsi logo. In the other two types of trials Pepsi would be administered, once cued with a Pepsi logo and once cued with a Coke logo. This would help to dissociate the role of brand information for preference encoding from memory functions linked to retrieving the brand associations.

A similar version of the above-suggested experiment has been done using a different methodological approach, namely using patients with brain damage or lesions in a specific brain area, here the vmPFC. Koenigs and Tranel (2008) investigated how preferences for Coke vs. Pepsi in patients with damage in the vmPFC changed during blind vs. open tasting of both sodas. They found that brand associations in the open tasting did not influence the lesion patients, only the control patients. In other words, patients with a lesioned vmPFC did not reverse their preferences when they knew what brand of soda they were consuming. To better understand the roles of the regions reported by McClure et al. (2004), similar studies should be conducted on patients with injury to the dorsolateral PFC and hippocampus. The advantage of using lesion patients as compared to fMRI is that causal and not “only” correlational links between mental processes and brain functioning can be established.

Intensity Another, much smaller stream of research has investigated the intensity of emotional and sensory experiences: In humans, subjective reports of pain intensity correlated with

activity in the insula and the ACC (Davis, Taylor, Crawley, Wood, & Mikulis, 1997; Peyron et al., 1999). Recent studies in the chemosensory domain found that amygdala activity increased with the intensity of negative and positive chemosensory stimuli (Anderson et al., 2003; Small et al., 2003). Several studies by Berns and colleagues suggest that the saliency or intensity of objects such as sound and money correlate with neural activity in the dorsal and ventral striatum (Zink, Pagnoni, Chappelow, Martin-Skurski, & Berns, 2006; Zink, Pagnoni, Martin, Dhamala, & Berns, 2003; Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004). Similar results have been found for the neural correlates of flavor intensity vs. flavor pleasantness (Small et al., 2003). To the best of our knowledge no studies to date have investigated how brand associations influence the intensity of experienced value signals. In turn, there are a lot of questions open for future research in this area.

Motivational value A concept that is related to how predicted and experienced values interact is the motivational value or incentive salience of an option. Over the past two decades, pioneering work by Berridge and colleagues has contributed to a better understanding of value processing in the brain by distinguishing between “wanting” and “liking” responses to stimuli (Berridge, 2007, 2009a, 2009b; Berridge & Kringelbach, 2008; Berridge & Robinson, 1998). “Wanting” refers to a person’s (or animal’s) motivation to obtain a given reward, as observed by increased effort, longer viewing times, and stronger grip strength (e.g., Pessiglione et al., 2007). “Liking” refers to the experienced value. This line of research has found an important role of the dopaminergic system for wanting, but not necessarily for liking (Berridge & Kringelbach, 2008; Berridge & Robinson, 1998).

A study by Litt, Plassmann, Shiv, and Rangel (2011) showed that the predicted value signals encoded in the mOFC/vmPFC are not confounded with related saliency signals of the options for choice.

More recently, the distinction between “wanting” and “liking” has also received more attention in consumer behavior research (Brendl, Markman, & Messner, 2003; Litt, Khan, & Shiv, 2010; Morewedge, Huh, & Vosgerau, 2010; Xianchi, Brendl, & Ariely, 2010). However, to date there is no behavioral or neuroscientific research trying to understand how different types of brand associations, and how favorable or unique they are, affect wanting and liking for brands.

Remembered value and learning

Consider again our example of choosing between Heineken and Beck’s. An important predictor of your choice is your memory of previous exposures to the two brands. If you remember that Heineken had a bitter taste and Beck’s had a distinct whiff of something sweet, these experiences may influence your decision. You may remember a recent entertaining Heineken commercial, but you have no such memory of a Beck’s commercial. Brands “work their magic” by associating themselves with experiences, which in turn influence subsequent retrieval and recognition. It is important to note that these can be personal experiences or those of other people—

shown in commercials or told by friends through word of mouth.

Remembered value refers to how different brand associations are encoded, consolidated, and retrieved in the consumer’s memory. Recent research suggests that parts of these processes happen on an unconscious level. Similar models have been seen in consumer psychology. For example, Van Osselaer and Janiszewski (2001) distinguished between the Human Associative Memory model, a process that was a general and unfocused incidental (or unconscious) associative learning, and an adaptive learning mechanism focusing on feature-benefit associations for future rewards.

Hence, the remembered value consists of both explicit memory and implicit memory of prior consumption experience. In this section, we first review the neuroscientific literature of explicit and implicit memory and learning processes relevant to branding. Following this, we briefly review recent studies on the dynamic nature of brand memories, that is, how explicit and implicit memory changes over time and how external factors such as marketing actions might affect remembered value.

Explicit brand memory Studies have demonstrated that explicit memories—also known as declarative memories—rely on specific brain regions such as the hippocampus and surrounding medial temporal lobe (MTL) region, in synchrony with other brain regions such as the dlPFC (Squire & Zola, 1996a, 1996b, 1998). Indeed, the distinction between declarative and non-declarative memories remains a dominant model for our understanding of memory function (Ramsøy et al., 2009; but see Henke, 2010, for a recent alternative account). Several studies have reported a strong link between memory and preference. In a seminal paper by McClure et al. (2004), it was reported that an increase in preference for the beverage labeled as Coca-Cola, but not the one labeled Pepsi Cola, was paralleled by an activation increase in the hippocampus and the dlPFC. In other words, the brand-induced change in preference was mediated by regions implicated in declarative memory.

Similar activations of the PFC were reported in the aforementioned study by Schaefer, Berens, Heinze, and Rotte (2006) in which subjects were asked to visually imagine driving a car of a well-known car manufacturer (e.g., BMW) or an unknown generic car brand in the German car market at the time the study was conducted (e.g., Acura). That is, while the act of imagining driving a car was equal between the two conditions, imagining driving a well-known car led to a stronger engagement of the superior frontal gyrus of the PFC, which has also been implicated in memory function.

In a recent study by Klucharev, Smidts, and Fernandez (2008) the link between memory and preference was further strengthened by studying how “expert power” influences this link. In the study, products that were presented simultaneously with an expert person were associated with improved recall at a subsequent memory test on a different day. Notably, by using fMRI during expert object presentations, the researchers found activation changes related to successful encoding and subsequent recall. Expert conditions were associated with increased activity in the striatum (both caudate

and likely the ventral striatum/NAcc) and, interestingly, with memory-related activity in the dlPFC, hippocampus, and parahippocampal cortex. Probing the relationship between favorable attitude toward experts and memory performance, the researchers also demonstrated a more direct coactivation of the bilateral caudate nuclei, hippocampus, and parahippocampal cortex. Thus, the link between preference and memory seems to be based on a synergic coactivation of the reward system and memory-related structures such as the dlPFC, hippocampus, and parahippocampal cortex. In another study, Schaefer and Rotte (2007b) demonstrated that brand names or logos engaged the neural reward system, further suggesting that a brand can work as a secondary reinforcer and act on the valuation systems of the brain.

Nevertheless, much is still unknown about the relationship between explicit memory and preference formation, and studies have shown inconsistencies. In another study by Schaefer and Rotte (2007a) it was found that when subjects viewed their most beloved brands, there was a decrease in the activation of both the dlPFC and the hippocampus and an increase in activation in reward regions such as the striatum. Thus, this study may seem at odds with previous suggestions of a positive relationship between memory engagement and preference formation. However, one may contend that the increased activation of memory-related regions found by Schaefer et al. (2006) can be explained by the greater visualization richness when imagining driving a well-known car brand compared to an unknown, generic car. Thus, this complicating factor may be related to differences in study design and other factors, but nevertheless highlights that the neural bases and the basic mechanisms of branding are still poorly understood.

Implicit brand memory As shown by several recent reports during the past few decades, the search for unconscious processes and implicit measures of branding is an active field of inquiry in consumer psychology (Baker, 2003; Bargh, 2002; Brasel and Gips, 2011; Chartrand et al., 2008; Claudi, Dimofte and Yalch, 2011; Friese, Wänke, & Plessner, 2006; Janiszewski, 1993; Moore, 1988; Nevid, 2010; Pratkanis & Greenwald, 1988; Saegert, 1987; Shapiro, 1999; Synodinos, 1988; Theus, 1994; Zajonc & Markus, 1985; Zaltman, 2000). For example Chartrand et al. (2008) demonstrated that subliminally presented retail brand names had an influence on goal pursuit. This suggests that the motivational effect of brands has an unconscious basis.

Although some scholars suggest a more cautious take on the power of the unconscious in the consumer research domain (see Simonson, 2005), recent insights from both behavioral approaches and neuroimaging make it inescapable that brands can be triggered unconsciously or, even when presented overtly, can affect consumer behavior without the person being privy to such effects.

An ongoing debate in cognitive neuroscience concerns the degree to which unconscious stimuli can affect processing in the brain and influence behaviors (Kouider & Dehaene, 2007). Most accounts postulate that low-level computations (e.g., motor reflexes and sensory processing) are driven by

unconscious mechanisms, while high-level executive functions such as decision making require consciousness. Yet an increasing body of evidence suggests that higher processing levels can be engaged unconsciously.

Recent studies have demonstrated that the prefrontal cortex, associated with conscious executive functions, can be engaged by subliminal task-switching cues (Lau & Passingham, 2007). For example, Pessiglione et al. (2007) reported that subliminal high-value rewards increased the strength with which subjects deployed effort on a hand grip task relative to low-value rewards. This was related to activation of the ventral striatum (VS), a subcortical reward structure.

Similar findings were made for abstract icons as primes, showing that reward-related activation and learning mechanisms of the VS could operate unconsciously (Pessiglione et al., 2008), and such findings suggest a role for unconscious learning processes in guiding motivated behaviors. From early studies of primates and mammals (Hollerman & Schultz, 1998; Schultz, 1998, 2001; Schultz & Dickinson, 2000) to more recent neuroimaging studies (Brown & Braver, 2007; D'Ardenne, McClure, Nystrom, & Cohen, 2008), value-based learning is now thought to include sub-cortical, low-level brain regions such as the ventral tegmental area, striatum, anterior cingulate cortex, and hippocampus.

Taken together, the implicit brand memories seem to engage both the deeper basic structures of the brain and memory regions previously thought to be dedicated solely to explicit memories. By using new and sophisticated analysis tools on neuroimaging data, it is possible to track neural processes that precede and even predict conscious choice. Such advances not only improve our understanding of implicit brand memory but provide a whole new avenue for studying the consumer psychology of branding.

The dynamic nature of memories Different models of memory retrieval have seen memories as being “replayed,” contributing to the popular notion that (episodic) memories are stored as hard copies in the brain. In this view, remembering is the process of retrieving factual and true information about the experiences we have had. While memories have been thought of as labile during encoding, information that is consolidated in memory has been thought to be retrieved as more stable “information packages.” Neurobiological studies have recently challenged this notion (Nader, Schafe, & LeDoux, 2000; Schafe, Nader, Blair, & LeDoux, 2001), showing that even powerful memories such as fear conditioning can be altered or eradicated by inhibiting the neural mechanisms (protein synthesis in the amygdala) during retrieval. This suggests that the retrieval stage is an active and dynamic relearning process rather than the mere replay of previously acquired information.

The idea of false memories has been paralleled by research in consumer psychology and behavior. A study by Braun-Latour and Zaltman (2006) demonstrated that advertising can unconsciously alter consumers’ beliefs as reflected by a change in how they recalled their earlier reporting of these beliefs following exposure to advertising. A related study by Cowley (2007) also showed that affective reactions derived from post-experience

information (i.e., advertising) may interfere with the retrieval of experience-based reactions: The results of three experiments showed that when post-experience affective reactions interfere with the retrieval of an experience-based reaction, consumers use post-experience behavior as a proxy for their liking of the experience. In a recent study making use of the dynamic nature of memory, [Rajagopal and Montgomery \(2011\)](#) demonstrated that exposure to an imagery-evoking ad led to a false memory of prior product exposure, further causing alterations in product attitude. Future studies should connect the consumer psychology and neuroscience literatures, and focus on the role and neural bases of dynamic memory in the formation, sustaining, and alteration of brand preferences.

Taken together, this review of “what has been done” shows that interdisciplinary effort in understanding how decision making is represented in the brain has taken off and that these findings can be applied to extend our understanding of the psychology of branding. Our review also points out directions for future research in this area. Two critical issues are important to note:

First, most of the studies reviewed above are mostly exploratory in nature, but have already succeeded in challenging our notions of how branding works. We encourage researchers to go beyond a mere correlation approach, i.e., localizing the neural bases of brand familiarity or brand preference. This type of research is important but contributes mostly to our understanding of the brain while providing fewer novel insights into the psychology of brands. To advance the psychology of branding, we suggest that future work focus more on establishing meaningful brain-behavior relationships that go beyond correlational findings, by combining the neuroscientific tool kit with traditional methods. We make suggestions to address this point in the last section of this paper.

Second, the review in this section pointed out several methodological issues with previous consumer neuroscience studies related to branding that have been published in academic journals. Many additional issues can be raised for unpublished work in the form of commercial applications of neuroscience to the psychology of branding, which has become a business with almost exponential growth. In the next section of the paper we detail the major issues and make suggestions for how to overcome them.

What should not be done: the need of standards for neuroscience work published in marketing journals

How can neuroimaging be a valuable tool for branding researchers? The promise of having a method for opening the “black box” of consumers’ brains may seem like a dream come true for any academic or practitioner interested in branding and other areas of consumer behavior. However, as seen in the review above, one can identify at least one major issue that needs the attention of researchers applying neuroscience tools for branding questions and of reviewers of such work: how to overcome the problem of reverse inference.

One practice that has become common in consumer neuroscience studies in general, and those related to branding in

particular, is reverse inference, by which the engagement of a particular mental process is inferred from the activation of a particular brain region ([Poldrack, 2006](#)). We believe the deductive validity of such inferences is limited.

The inference that is usually drawn from neuroimaging data, according to current scientific standards, is of the form “if cognitive process X (e.g., willingness-to-pay computation) is engaged, then brain area Z (e.g., mOFC) is active.” However, previous consumer neuroscience and commercial neuromarketing studies reverse this reasoning, as follows:

- In the current study, when task comparison A was presented (e.g., imagining driving a car branded by a familiar vs. an unfamiliar logo seen on the screen), brain area Z (e.g., the medial prefrontal cortex) was active.
- In other studies, when cognitive process X (e.g., self-reflection and self-relevant thoughts) was putatively engaged, brain area Z (e.g., the medial prefrontal cortex) was active.
- Thus, the activity of area Z (e.g., the medial prefrontal cortex) in the current study demonstrates engagement of cognitive process X (e.g., self-reflection and self-relevant thoughts) by task A (e.g., imagining driving a car branded by a familiar vs. an unfamiliar logo seen on the screen).

This has been referred to as a reverse inference, since it reasons backwards from the presence of brain activation to the engagement of a particular mental process ([Poldrack, 2006](#)). The fact that reverse inference is problematic is partly due to the fact that functional brain imaging research is still relatively new, and as a consequence, we do not have a detailed map of the brain to date. More important is the fact that a single brain area can multitask, and that the brain has a built-in redundancy. In other words, one particular brain area could be involved in encoding both brand personality associations and brand familiarity. If a study finds this brain area Z to be involved in brand decisions without implementing a design that allows dissociating between the two, the inference that activity in this area means that one brand is more familiar is of only limited validity.

In many cases the use of reverse inference is informal; the presence of unexpected activation in a particular region is explained by reference to other studies that found activation in the same region. The issue of reverse inference becomes much more problematic when the central findings and contributions of the paper are built on reverse inference. As our review of consumer neuroscience studies related to branding revealed, several of the previous studies in this area (these authors’ included) use reverse inference as a central feature to discuss their findings. There are several ways to address the problem of reverse inference in neuroimaging studies.

The first and most straightforward way is to implement an experimental design and data analysis that allow capturing the neural signature of the mental process of interest directly. For example, a recent study investigated how changing the price of wine affects taste processing in the brain ([Plassmann et al., 2008](#)). The study found that when subjects consumed the same wine in two experimental conditions, once cued with a

high price and once with a low price, brain activity in the mOFC was affected. Other studies found that this brain area encoded taste pleasantness. The design of the study allowed the authors to run a different data model to check which brain area in their data encoded taste pleasantness. They found that in their data also the mOFC encoded taste pleasantness irrespective of the changes in the price of the wine. Their design and data analysis procedure allowed the authors to control for relying on reverse inference.

The second way to address the reverse inference problem is to find a measure of the degree to which the region of interest is selectively activated by the mental process of interest (Ariely & Berns, 2010; Poldrack, 2006). If, on one hand, a region is activated by a large number of mental processes, then activation in that region provides relatively weak evidence of the engagement of the mental process. If, on the other hand, the region is activated relatively selectively by the specific mental process of interest, then one can infer with substantial confidence that the process is engaged given activation in the region. The idea is to compute a selectivity factor that determines the posterior probability for the reverse inference using Bayesian statistics based on previous findings (see Poldrack, 2006 for details).

However, there are at least two important concerns. First, although Poldrack's procedure to compute a selectivity factor is meaningful in a statistical sense, the assumptions behind such a calculation are rather liberal and may suffer from a publication bias for positive results. Second, the mental process of interest needs to be specified very precisely for an application of this idea to consumer neuroscience. "Reward processing" seems rather general, and the question remains whether this refers to the prediction or the experience of reward. In other words, different and imprecise definitions of "reward" are problematic.

Taken together, the application of such a selectivity factor for judging whether reverse inference is possible needs to be done with caution. Given the limited power of reverse inference from single-region brain activations, more sophisticated multivariate methods for interpreting brain imaging data have been at the forefront of analysis techniques. The idea behind these techniques and how consumer neuroscience research related to branding may benefit from those is detailed in the next section of this paper.

What could be done: conclusions and suggestions for future directions

In this last section of the paper we lay out our vision of future consumer neuroscience research and why we think academics and practitioners alike could and should be excited about this new field. Since we have already provided concrete future directions for branding research in our review of what is currently done, we conclude with a broader view on the new directions the field of consumer neuroscience could take to make a substantial contribution to consumer research and the psychology of branding.

The application of neuroscience to consumer psychology, specifically to the psychology of branding, has an interesting

potential for at least two reasons. First, it can be viewed as a new methodological tool, as a "magnifying glass" to observe mental processes without asking consumers directly for their thoughts, memories, evaluations, or decision-making strategies, and thus can provide access to otherwise hidden information (Ariely & Berns, 2010; Plassmann et al., in press). Second, neuroscience can be viewed as a source of theory generation, supplementing traditional theories from psychology, marketing, and economics (Plassmann et al., in press). We explain both ideas in the following section.

Neuroscience as a tool

Neuroscience's potential as a tool stems from at least two ways it can contribute to a better understanding of the psychology underlying brands. First, combining advanced statistical models from computer science with neuroscience data makes it possible to predict behavior in a more accurate way than relying on traditional measures such as self-reports. Second, by combining different tools from the neuroscientific tool kit we can establish brain-behavior relationships that are meaningful for understanding the psychology underlying consumer choices.

Predicting consumer choices Empirical studies in consumer neuroscience and neuromarketing employ neuroimaging tools as biomarkers to assess responses to marketing stimuli such as brands, advertisements, packaging and to predict consumer choices.

For example, in a study by Knutson et al. (2007), subjects, while being scanned using fMRI, first saw the product (4 s), then were shown the price of the product (4 s), and finally made their choice (4 s). Subjects reported making their decision consciously only at the very end of each run (i.e., the last 4 s), yet analysis of the fMRI data showed the neural predictors of purchase at earlier time points. Notably, these activation changes could be traced from 8 to 12 s before the decision was made, and before subjects reported having made up their minds. However, the neural predictors did not demonstrate better predictive power than self-reports (pseudo- R^2 was 0.528 when only self-reports were included and changed to 0.533 when neural predictors were added; note that pseudo- R^2 was 0.105 based on neural predictors alone). Taken together, Knutson and colleagues could extract neural predictors at a time when subjects had not made up their minds yet, but these predictors were not fundamentally better at predicting purchase behavior than simply asking the subjects about their preferences.

Another example is a recent study by Berns and Moore (2012) that used a small group of subjects' neural responses to music to predict subsequent market level impact in form of commercial success of the songs (using sales data for a period of three years after the experiment). Interestingly, subjective liking ratings of songs did not correlate with future sales data, but the neural response did (i.e., brain activation within the nucleus accumbens).

New developments in neural pattern classification techniques and multivariate decoding analysis of fMRI data (Haynes & Rees, 2006) are very promising to increase the

predictive power of neuroscientific tools in the years to come. A first study in the context of consumer behavior was done by [Tusche, Bode, and Haynes \(2010\)](#). In their study subjects were presented with images of different cars, and asked either to rate their liking of each car (high-attention group) or perform a visual fixation task (low-attention group). After the task, subjects rated their willingness to buy each car. Crucially, subjects were scanned using fMRI during the task, allowing the researchers to test whether neural activation could predict subsequent car choice. The fMRI data were analyzed using a multivariate analysis approach, in which data were fed into the analysis, showing brain regions between the high- and low-attention groups that predicted subsequent purchase intentions.

The idea behind these techniques is as follows. Whole-brain neuroimaging data acquisition, such as fMRI, generates time series data from thousands of data points across the brain. While standard analyses of neuroimaging data employ large-scale univariate analyses by contrasting different experimental conditions, multivariate pattern classification techniques take advantage of information contained in multiple voxels distributed across space. They allow investigation of whether spatial patterns of brain activation contain stable information about different experimental conditions (e.g., purchase vs. no purchase).

These approaches promise better predictions of decision-making behavior across domains, such as neural, physiological, and behavioral predictors of in-store purchase, unhealthy behaviors, and overspending. We believe that decoding of brain patterns using such sophisticated algorithms will be a turning point for consumer neuroscience research.

Establishing brain–behavior relationships that are meaningful for consumer psychology Another potential way to apply methodologies from neuroscience to consumer behavior is to observe consumers' mental processes in real time. As discussed earlier, this is of particular importance when the underlying processes are difficult to investigate because they are below consumers' awareness or are difficult to verbalize and/or manipulate in a traditional experimental setting or survey. One example was provided by a recent study ([Plassmann et al., 2008](#)) that investigated whether marketing actions (i.e., changing the price of a wine) does alter taste processing (i.e., the wine actually tastes better) or cognitive processing because of rationalizing (i.e., the consumer thinks the wine tastes better). It is very difficult for consumers to verbalize whether the price changes how much they think they like the wine or how much they actually like the wine, although this difference is very important from a consumer psychology perspective. The authors could show that changing the price of an identical wine does actually change taste processing and more specifically that part of taste processing that encodes the pleasantness of the taste. This finding provides neuropsychological evidence for a placebo effect of marketing actions on positive experiences similar to placebo effects in the pain domain.

Another approach is to use neuroscientific measures to validate behavioral measures. An example of this approach is a

recent study ([Dietvorst et al., 2009](#)) that aimed at developing a sales force-specific Theory-of-Mind (ToM) scale in two steps. First, they developed a personality scale measuring salespeople's interpersonal-mentalizing skills, based on questionnaires. Second, they validated the questionnaire-based scale by comparing high- and low-scoring salespeople on the scale when they worked on interpersonal-mentalizing and control tasks while having their brains scanned using fMRI. Interestingly, they found that salespeople who scored high on their sales force-specific ToM scale also showed more activation in brain areas involved in ToM during the interpersonal-mentalizing tasks but not during the control tasks.

It is important to note that the next level of research in this area needs to go beyond merely establishing associations between brain activity and a specific behavior. A review by [Kable \(2011\)](#) showed that 60% to 70% of empirical studies applying neuroscience to behavioral decision-making theories use only one method: fMRI. To establish a deeper understanding of the relationships between neuropsychological processes and behavior that can profoundly advance our understanding of consumer psychology, consumer neuroscientists need to expand the neuroscientific tool kit. The idea behind this is to show that (a) brain mechanisms are necessary for a specific consumer behavior (i.e., when brain activity is interrupted, behavior is impaired) and (b) brain mechanisms are sufficient for a specific consumer behavior (i.e., when brain activity is induced, behavioral effects occur; see [Kable, 2011](#), for a more detailed discussion).

Methods to test necessity include using patients who have a lesion in a specific brain area of interest, such as the vmPFC, and testing their behavior as compared to control populations. For example, it has been shown that focal brain lesions in this area make patients outperform healthy controls in financial performance tasks ([Shiv, Loewenstein, & Bechara, 2005](#)). Another way to study necessity is the application of techniques that "virtually lesion" healthy subjects by temporarily interrupting electromagnetic activity (Transcranial Magnetic Stimulation (TMS) or cathodal Transcranial Direct Current Stimulation (TDCS)). A recent study by [Camus et al. \(2009\)](#) showed that the application of inhibitory TMS to subjects' dlPFC decreased subjects' predicted values in an economic auction.

The toolkit to test sufficiency is much more limited and includes primarily a reversed version of transcranial direct current stimulation (anodal TDCS). For example, a study by [Fregni et al. \(2008\)](#) showed that stimulation of the lateral prefrontal cortex reduced craving in smokers.

Beyond testing relationships between brain systems and behavior, another novel and exciting approach is to go one level deeper and test the relationships between specific neurotransmitters and behavior. Recent advances in our understanding of the role of neurotransmitters, and how they relate to processes underlying decision making, may lead to improved understanding of consumer psychology. Few studies, if any, have approached this from a consumer behavior perspective, but insights from studying decision-making on a neurotransmitter-level might serve as a source to generate new research ideas (see [Ramsøy & Skov, 2010](#), for a review). Applying the same

idea described above, the neuroscientific tool kit allows us to test associations, necessity, and sufficiency of neurotransmitters and specific consumer behavior (see Kable, 2011, for a more detailed discussion).

Specific brain imaging techniques that allow tracking of changes in neurotransmitters (forms of Positron Emission Tomography (PET)) and the study of genetics allow researchers to make associations between neurotransmitters such as dopamine and a specific behavior such as gambling or other impulsive behaviors.

Administration of pharmacological antagonists or depletion of a specific neurotransmitter (e.g., through dietary restrictions) allows researchers to test necessity. For example, a study by Crockett, Clark, Tabibnia, Lieberman, and Robbins (2008) found that serotonin depletion increased rejection of unfair offers in an ultimatum game.

Along those lines, administration of pharmacological agonist or depletion of a specific neurotransmitter allows researchers to test *specificity*. For example, Kosfeld, Heinrichs, Zak, Fischbacher, and Fehr (2005) demonstrated that administration of oxytocin increased trust during economic exchange. Another example is a study by Schweighofer et al. (2008), who tested the effect of serotonin loading and depletion on reward discounting.

Taken together, studies in consumer psychology can benefit from new tools that allow the testing of association, necessity, and sufficiency of neuropsychological processes and consumer behavior. By expanding the toolbox in consumer neuroscience, advances can be made in our understanding of both basic mechanisms and individual differences in consumer decision making.

Neuroscience as basis for theory generation

Although most of the hype around the potential of consumer neuroscience and neuromarketing evolves around using neuroscientific tools, in this review we would like to suggest neuroscientific findings as a novel source of understanding the mechanisms underlying consumer psychology, as pioneered by Wadhwa, Shiv, and Nowlis (2008) and others (e.g., Lee, Amir, & Ariely, 2009; Van Den Bergh, Dewitte, & Warlop, 2008).

Wadhwa and colleagues investigated the effect of product sampling at the point of sale on subsequent consumption behavior (Wadhwa et al., 2008). The authors compared different hypotheses about whether product sampling would increase subsequent consumption behavior, and if so, whether the effects would be specific to the product sampled, to its product category, or to anything perceived as pleasurable. These predictions were based on different theories from psychology, physiology, and neurophysiology of taste and reward. In a series of experiments, the authors found support for the prediction that our general motivation system in the brain is at work when we sample products, leading to an increased subsequent reward-seeking behavior for any other type of reward. Similarly, a study from Van Den Bergh et al. (2008) found impatience in intertemporal choice

to be linked to the activation of the general motivation system in the brain.

Another example is a recent study by Ramsøy, Loving, Skov, and Clement (2011) in which women were studied during different phases of their ovarian cycle. It is well known that this cycle has significant effects on female thinking and behavior, including changes in memory, sexual behavior, and mate selection (Jones et al., 2008; Pillsworth, Haselton, & Buss, 2004; Rupp & Wallen, 2007; Vranić & Hromatko, 2008; Zhu et al., 2010). In particular, recent studies have demonstrated changes in consumer behavior, including the increased likelihood of purchasing and wearing sexually suggestive clothing at peak fertility (Durante, Li, & Haselton, 2008; Durante et al., 2010), although Saad and Stenstrom (2012), interestingly, did not find evidence linking menstrual cycle to attitudes towards brand-related information. Little is known about the exact mechanisms underlying such effects and to what extent menstrual cycle affects the processing of different kinds of brands or advertisements. By using eye tracking to assess visual attention, Ramsøy et al. (2011) found that at peak fertility, women tended to show faster and more frequent fixations and longer total viewing time toward sexual elements in ads. Such effects were not at the cost of visual attention toward brand information and did not have an impact on preference or long-term memory scores. Nevertheless, these findings demonstrate how a known biological factor may influence consumer psychology.

These studies are examples of how scholars in consumer psychology can integrate findings and concepts from neuroscience without actually applying neuroscientific methods. This approach is of great potential for developing an interdisciplinary understanding of how consumers make decisions and may provide significant improvements in our understanding of preference formation and decision making. We hope this review will help researchers as a starting point for generating hypotheses based on an interdisciplinary framework to advance existing theories in consumer psychology.

To conclude, in this last section of this critical review, we have pointed out two major new directions in which neuroscience might advance consumer psychology. These new directions extend first findings in the nascent field of consumer neuroscience related to branding and, more important, help to address the issues of previous work reviewed in this paper. We hope this review provides researchers with exciting new perspectives and ideas for their future work in consumer neuroscience to advance our understanding of the psychology of branding.

Acknowledgments

The authors would like to thank Beth Pavlicek, Anupama Deepa Kumar, and the participants of the INSEAD course “Decision-Making and the Human Brain” for their comments on previous versions of this paper. We thank the participants of the tutorial “Decision Neuroscience” at the University of California at San Diego and University of Southern California for their insightful comments on aspects addressed in this paper. We also

thank the editors and the two anonymous reviewers for their valuable feedback and comments.

References

- Aaker, J. L. (1997). Dimensions of brand personality. *Journal of Marketing Research*, 34, 347–356.
- Aaker, J., & Fournier, S. (1995). A brand as a character, a partner, and a person: Three perspectives on the question of brand personality. *Advances in Consumer Research*, 22, 391–395.
- Aarts, H., Dijksterhuis, A., & De Vries, P. (2001). On the psychology of drinking: Being thirsty and perceptually ready. *British Journal of Psychology*, 92(4), 631–642.
- Aggarwal, P. (2004). The effects of brand relationship norms on consumer attitudes and behavior. *Journal of Consumer Research*, 31, 87–101.
- Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, 32(3), 537–551.
- Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6(2), 196–202.
- Ariely, D., & Berns, G. S. (2010). Neuromarketing: The hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, 11(4), 284–292.
- Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biassing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, 3, 396–403.
- Baker, W. E. (2003). Does brand name imprinting in memory increase brand information retention? *Psychology and Marketing*, 20(12), 1119–1135.
- Ballard, K., & Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage*, 45(1), 143–150.
- Bargh, J. A. (2002). Losing consciousness: Automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research*, 29(2), 280–285.
- Berns, G., & Moore, S. E. (2012). A neural predictor of cultural popularity. *Journal of Consumer Psychology*, 22, 154–160.
- Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology (Berl)*, 191(3), 391–431.
- Berridge, K. C. (2009). "Liking" and "wanting" food rewards: Brain substrates and roles in eating disorders. *Psychology & Behavior*, 97(5), 537–550.
- Berridge, K. C. (2009). Wanting and liking: Observations from the neuroscience and psychology laboratory. *Inquiry (Oslo)*, 52(4), 378.
- Berridge, K. C., & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, 199(3), 457–480.
- Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, 28(3), 309–369.
- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences (USA)*, 98(20), 11818–11823.
- Brasel, S. A., & Gips, J. (2011). Red Bull "Gives You Wings" for better or worse: A double-edged impact of brand exposure on consumer performance. *Journal of Consumer Psychology*, 21(1), 57–64.
- Braun-Latour, K. A., & Zaltman, G. (2006). Memory change: An intimate measure of persuasion. *Journal of Advertising Research*, 46(1), 57–72.
- Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30(2), 619–639.
- Brendl, C. M., Markman, A. B., & Messner, C. (2003). The devaluation effect: Activating a need devalues unrelated objects. *Journal of Consumer Research*, 29(4), 463–473.
- Brown, J. W., & Braver, T. S. (2007). Risk prediction and aversion by anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 266–277.
- Camus, M., Halelamien, N., Plassmann, H., Shimojo, S., O'Doherty, J., Camerer, C., & Rangel, A. (2009). Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex decreases valuations during food choices. *European Journal of Neuroscience*, 30(10), 1980–1988.
- Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, 73(6), 1–17.
- Chartrand, T. L., Huber, J., Shiv, B., & Tanner, R. J. (2008). Nonconscious goals and consumer choice. *Journal of Consumer Research*, 35(2), 189–201.
- Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39), 12315–12320.
- Cowley, E. (2007). How enjoyable was it? Remembering an affective reaction to a previous consumption experience. *Journal of Consumer Research*, 34(4), 494–505.
- Crockett, M. J., Clark, L., Tabibnia, G., Lieberman, M. D., & Robbins, T. W. (2008). Serotonin modulates behavioral reactions to unfairness. *Science*, 320(5884), 1739.
- D'Ardenne, K., McClure, S. M., Nystrom, L. E., & Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*, 319(5867), 1264–1267.
- Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology. *International Journal of Personality and Social Psychology*, 58(2), 330–341.
- Davis, K. D., Taylor, S. J., Crawley, A. P., Wood, M. L., & Mikulis, D. J. (1997). Functional MRI of pain- and attention-related activations in the human cingulate cortex. *Journal of Neurophysiology*, 77(6), 3370–3380.
- de Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C., & Cayeux, I. (2005). Cognitive modulation of olfactory processing. *Neuron*, 46(4), 671–679.
- Deppe, M., Schwindt, W., Kramer, J., Kugel, H., Plassmann, H., Kenning, P., & Ringelstein, E. B. (2005). Evidence for a neural correlate of a framing effect: Bias-specific activity in the ventromedial prefrontal cortex during credibility judgments. *Brain Research Bulletin*, 67, 413–421.
- Deppe, M., Schwindt, W., Kugel, H., Plassmann, H., & Kenning, P. (2005). Non-linear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *Journal of Neuroimaging*, 15, 171–182.
- Deppe, M., Schwindt, W., Pieper, A., Kugel, H., Plassmann, H., Kenning, P., Deppe, K., & Ringelstein, E. B. (2007). Anterior cingulate reflects susceptibility to framing during attractiveness evaluation. *Neuroreport*, 18, 1119–1123.
- Dietvorst, R. C., Verbeke, W. J. M. I., Bagozzi, R. P., Yoon, C., Smits, M., & van der Lugt, A. (2009). A sales force-specific theory-of-mind scale: Tests of its validity by classical methods and functional magnetic resonance imaging. *Journal of Marketing Research*, 46(5), 653–668.
- Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology*, 61, 467–490.
- Dimofte, C. V., & Yalch, R. F. (2011). The mere association effect and brand evaluations. *Journal of Consumer Psychology*, 21(1), 24–37.
- Dreze, X., & Hussell, F. -X. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing*, 17(4), 8–23.
- Durante, K. M., Griskevicius, V., Hill, S. E., Perilloux, C., Li, N. P., & Nordqvist, C. (2010). Ovulation, female competition, and product choice: Hormonal influences on consumer behavior. *Journal of Consumer Research*, 37, 921–934.
- Durante, K. M., Li, N. P., & Haselton, M. G. (2008). Changes in women's choice of dress across the ovulatory cycle: Naturalistic and laboratory task-based evidence. *Personality and Social Psychology Bulletin*, 34(11), 1451–1460.
- Durgin, F. H., Doyle, E., & Egan, L. (2008). Upper-left gaze bias reveals competing search strategies in a reverse Stroop task. *Acta Psychologica*, 127(2), 428–448.
- Efron, R., & Yund, E. W. (1996). Spatial nonuniformities in visual search. *Brain and Cognition*, 31(3), 331–368.
- Egidi, G., Nusbaum, H. C., & Cacioppo, J. T. (2008). *Neuroeconomics: Foundational issues and consumer relevance*. Mahwah, NJ: Erlbaum.

- Erk, S., Spitzer, M., Wunderlich, A. P., Galley, L., & Walter, H. (2002). Cultural objects modulate reward circuitry. *Neuroreport*, 13, 2499–2503.
- Esch, F. R., Möll, T., Schmitt, B., Elger, C. E., Neuhaus, C., & Weber, B. (2012). Brands on the brain: What happens neurophysiologically when consumers process and evaluate brands? *Journal of Consumer Psychology*, 22, 75–85.
- Fournier, S. (1997). Consumers and their brands: Developing relationship theory in consumer research. *Journal of Consumer Research*, 24(4), 343–373.
- Fregni, F., Liguori, P., Fecteau, S., Nitsche, M. A., Pascual-Leone, A., & Boggio, P.S. (2008). Cortical stimulation of the prefrontal cortex with transcranial direct current stimulation reduces cue-provoked smoking craving: A randomized, sham-controlled study. *Journal of Clinical Psychiatry*, 69, 32–40.
- Friesé, M., Wänke, M., & Plessner, H. (2006). Implicit consumer preferences and their influence on product choice. *Psychology and Marketing*, 23(9), 727–740.
- Glaholt, M. G., & Reingold, E. M. (2009). Stimulus exposure and gaze bias: A further test of the gaze cascade model. *Attention, Perception, & Psychophysics*, 71(3), 445–450.
- Glaholt, M. G., Wu, M. C., & Reingold, E. M. (2010). Evidence for top-down control of eye movements during visual decision making. *Journal of Vision*, 10(5), 15.
- Grabenhorst, F., Rolls, E. T., & Bilderbeck, A. (2008). How cognition modulates affective responses to taste and flavor: Top-down influences on the orbitofrontal and pregenual cingulate cortices. *Cerebral Cortex*, 18(7), 1549–1559.
- Habib, R., Nyberg, L., & Tulving, E. (2003). Hemispheric asymmetries of memory: The HERA model revisited. *Trends in Cognitive Sciences*, 7(6), 241–245.
- Hare, T. A., Camerer, C. F., Knoepfle, D. T., & Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *Journal of Neuroscience*, 30(2), 583–590.
- Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *Journal of Neuroscience*, 28(22), 5623–5630.
- Harmon-Jones, E. (2003). Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology*, 40(6), 838–848.
- Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.
- Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience*, 11(7), 523–532.
- Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, 1(4), 304–309.
- Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11), 1254–1259.
- Janiszewski, C. (1993). Preattentive mere exposure effects. *Journal of Consumer Research*, 20(3), 376–392.
- Jones, B. C., DeBruine, L. M., Perrett, D. I., Little, A. C., Feinberg, D. R., & Law Smith, M. J. (2008). Effects of menstrual cycle phase on face preferences. *Archives of Sexual Behavior*, 37(1), 78–84.
- Junghofer, M., Kissler, J., Schupp, H. T., Putsche, C., Elling, L., & Dobel, C. (2010). A fast neural signature of motivated attention to consumer goods separates the sexes. *Frontiers in Human Neuroscience*, 4, 179.
- Kable, J. W. (2011). The cognitive neuroscience toolkit for the neuroeconomist: A functional overview. *Journal of Neuroscience, Psychology, and Economics*, 4(2), 63–84.
- Kahneman, D., & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making*, 5(3), 187–200.
- Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, 112(2), 375–405.
- Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, 57(1), 1–22.
- Kenning, P. H., & Plassmann, H. (2008). How neuroscience can inform consumer research. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(6), 532–538.
- Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fMRI study. *Neuroimage*, 44(3), 1125–1132.
- Klucharev, V., Smidts, A., & Fernandez, G. (2008). Brain mechanisms of persuasion: How “expert power” modulates memory and attitudes. *Social Cognitive and Affective Neuroscience*, 3(4), 353–366.
- Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, 30, 57–78.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectivity recruits nucleus accumbens. *Journal of Neuroscience*, 21(16), RC159.
- Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology*, 18(4), 411–417.
- Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L., & Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*, 12(17), 3683–3687.
- Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fMRI. *Neuroimage*, 18(2), 263–272.
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, 53(1), 147–156.
- Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., & Loewenstein, G. (2008). Neural antecedents of the endowment effect. *Neuron*, 58(5), 814–822.
- Koch, C. (2004). *Quest for consciousness: A neurobiological approach*. Englewood, CO: Roberts & Company Publishers.
- Koenigs, M., & Tranel, D. (2008). Prefrontal cortex damage abolishes brand-cued changes in cola preference. *Social Cognitive and Affective Neuroscience*, 3, 1–6.
- Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042), 673–676.
- Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society B*, 362(1481), 857–875.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Kringelbach, M. L., O'Doherty, J., Rolls, E. T., & Andrews, C. (2003). Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex*, 13(10), 1064–1071.
- Lau, H. C., & Passingham, R. E. (2007). Unconscious activation of the cognitive control system in the human prefrontal cortex. *Journal of Neuroscience*, 27(21), 5805–5811.
- Lee, L., Amir, O. N., & Ariely, D. (2009). In search of *Homo economicus*: Cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research*, 36(2), 173–187.
- Lee, N., Broderick, A. J., & Chamberlain, L. (2007). What is “neuromarketing”? A discussion and agenda for future research. *International Journal of Psychophysiology*, 63(2), 199–204.
- Leven, W. (1991). *Blickverhalten von Konsumenten. Grundlagen, Messung und Anwendung in der Werbeforschung*. Heidelberg: Physica Verlag.
- Litt, A., Khan, U., & Shiv, B. (2010). Lusting while loathing. *Psychological Science*, 21(1), 118–125.
- Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex*, 21, 95–102.
- Lohse, G. L. (1997). Consumer eye movement patterns on yellow page advertising. *Journal of Advertising*, 26, 61–73.
- McCabe, C., Rolls, E. T., Bilderbeck, A., & McGlone, F. (2008). Cognitive influences on the affective representation of touch and the sight of touch in the human brain. *Social Cognitive and Affective Neuroscience*, 3(2), 97–108.
- McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44(2), 379–387.
- Milosavljevic, M. (2009). *Computational modeling of visual attention and consumer research: Initial allocation of visual attention and its effects on consumer behavior*. Saarbruecken, Germany: VDM Verlag.
- Milosavljevic, M., & Cerf, M. (2008). First attention then intention: Insights from computational neuroscience of vision. *International Journal of Advertising*, 27(3), 381–398.
- Milosavljevic, M., Koch, C., & Rangel, A. (2011). Consumers can make choices in as little as a third of a second. *Judgment and Decision Making*, 6(6), 520–530.
- Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2011). Relative visual saliency differences induce sizable bias in consumer choice. *Working Paper*: California Institute of Technology.

- Moore, T. E. (1988). The case against subliminal manipulation. *Psychology and Marketing*, 5(4), 297–316.
- Morewedge, C. K., Huh, Y. E., & Vosgerau, J. (2010). Thought for food: Imagined consumption reduces actual consumption. *Science*, 330(6010), 1530–1533.
- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). The labile nature of consolidation theory. *Nature Reviews Neuroscience*, 1(3), 216–219.
- Nevid, J. S. (2010). Introduction to the special issue: Implicit measures of consumer response—The search for the holy grail of marketing research. *Psychology and Marketing*, 27(10), 913–920.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Norton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science*, 171(3968), 308–311.
- O'Doherty, J., Kringlebach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4(1), 95–102.
- O'Doherty, J., Rolls, E. T., Francis, S., Bowtell, R., McGlone, F., Kobal, G., Renner, B., & Ahne, G. (2000). Sensory-specific satiation-related olfactory activation of the human orbitofrontal cortex. *Neuroreport*, 11(4), 893–897.
- Ohme, R., Reykowska, D., Wiener, D., & Choromanska, A. (2009). Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures. *Journal of Neuroscience, Psychology, and Economics*, 2, 21–31.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge and New York: Cambridge University Press.
- Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, 59(4), 561–567.
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R. J., & Frith, C. D. (2007). How the brain translates money into force: A neuroimaging study of subliminal motivation. *Science*, 316(5826), 904–906.
- Peyron, R., Garcia-Larrea, L., Gregoire, M. C., Costes, N., Convers, P., Lavenne, F., Mauguire, F., Michel, D., & Laurent, B. (1999). Haemodynamic brain responses to acute pain in humans: Sensory and attentional networks. *Brain*, 122(9), 1765–1780.
- Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing*, 16(1), 1–16.
- Pieters, R., & Wedel, M. (2007). Goal control of visual attention to advertising: The Yarbus implication. *Journal of Consumer Research*, 34, 224–233 (August).
- Pillsworth, E. G., Haselton, M. G., & Buss, D. M. (2004). Ovulatory shifts in female sexual desire. *Journal of Sex Research*, 41(1), 55–65.
- Plassmann, H., Ambler, T., Braeutigam, S., & Kenning, P. (2007). What can advertisers learn from neuroscience? *International Journal of Advertising*, 26(2), 151–175.
- Plassmann, H., Kenning, P., & Ahlert, D. (2007). Why companies should make their customers happy: The neural correlates of customer loyalty. *Advances in Consumer Research—North American Conference Proceedings*, 34, (pp. 735–739).
- Plassmann, H., & Niessing, J. (2010). Expectation biases as neuropsychological basis for branding. German original title: Expectation Biases als neuropsychologische Grundlage des Markenmanagements. In M. Bruhn, & R. Köhler (Eds.), *Impulse aus der Neuroökonomie für die Markenführung*, 119–130. Wiesbaden: Gabler.
- Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, 27(37), 9984–9988.
- Plassmann, H., O'Doherty, J. P., & Rangel, A. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *Journal of Neuroscience*, 30(32), 10799–10808.
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences (USA)*, 105(3), 1050–1054.
- Plassmann, H., Yoon, C., Feinberg, F., & Shiv, B. (in press). Consumer neuroscience. In R. P. Bagozzi & A. Ruvio (Eds.), *Wiley International Encyclopedia of Marketing*. West Sussex, UK: John Wiley & Sons.
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10(2), 59–63.
- Pratkanis, A. R., & Greenwald, A. G. (1988). Recent perspectives on unconscious processing: Still no marketing applications. *Psychology and Marketing*, 5(4), 337–353.
- Rajagopal, P., & Montgomery, N. V. (2011). I imagine, I experience, I like: The false experience effect. *Journal of Consumer Research*, 38.
- Ramsøy, T. Z., Liptrot, M. G., Skimminge, A., Lund, T. E., Sidars, K., Christensen, M. S., Baare, W., Paulson, O. B., & Jernigan, T. L. (2009). Regional activation of the human medial temporal lobe during intentional encoding of objects and positions. *Neuroimage*, 47(4), 1863–1872.
- Ramsøy, T. Z., Loving, P., Skov, M., & Clement, J. (2011). Ovarian cycle impacts on women's visual attention towards sex in advertising. *Working Paper*. : Copenhagen Business School.
- Ramsøy, T. Z., & Skov, M. (2010). How genes make up your mind: Individual biological differences and value-based decisions. *Journal of Economic Psychology*, 31(5), 818–831.
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7), 545–556.
- Reutskaya, E., Nagel, R., Camerer, C., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101, 900–926.
- Rolls, E. T., Grabenhorst, F., & Franco, L. (2009). Prediction of subjective affective state from brain activations. *Journal of Neurophysiology*, 101(3), 1294–1308.
- Rolls, E. T., & McCabe, C. (2007). Enhanced affective brain representations of chocolate in cravers vs. non-cravers. *European Journal of Neuroscience*, 26(4), 1067–1076.
- Rosbergen, E., Pieters, R., & Wedel, M. (1997). Visual attention to advertising: A segment-level analysis. *Journal of Consumer Research*, 24, 305–314.
- Rupp, H. A., & Wallen, K. (2007). Sex differences in viewing sexual stimuli: An eye-tracking study in men and women. *Hormones and Behavior*, 51(4), 524–533.
- Saad, G., & Stenstrom, E. (2012). Calories, beauty, and ovulation: The effects of the menstrual cycle on food and appearance-related consumption. *Journal of Consumer Psychology*, 22, 102–113.
- Saegert, J. (1987). Why marketing should quit giving subliminal advertising the benefit of the doubt. *Psychology and Marketing*, 4(2), 107–120.
- Schaefer, M., Berens, H., Heinze, H. J., & Rotte, M. (2006). Neural correlates of culturally familiar brands of car manufacturers. *Neuroimage*, 31, 861–865.
- Schaefer, M., & Rotte, M. (2007). Favorite brands as cultural objects modulate reward circuit. *Neuroreport*, 18(2), 141–145.
- Schaefer, M., & Rotte, M. (2007). Thinking on luxury or pragmatic brand products: Brain responses to different categories of culturally based brands. *Brain Research*, 1165, 98–104.
- Schafe, G. E., Nader, K., Blair, H. T., & LeDoux, J. E. (2001). Memory consolidation of Pavlovian fear conditioning: A cellular and molecular perspective. *Trends in Neurosciences*, 24(9), 540–546.
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80(1), 1–27.
- Schultz, W. (2001). Reward signaling by dopamine neurons. *Neuroscientist*, 7(4), 293–302.
- Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Reviews of Neuroscience*, 23, 473–500.
- Schweighofer, N., Bertin, M., Shishida, K., Okamoto, Y., Tanaka, S. C., Yamawaki, S., & Doya, K. (2008). Low-serotonin levels increase delayed reward discounting in humans. *Journal of Neuroscience*, 28, 4528–4532.
- Shapiro, S. (1999). When an ad's influence is beyond our conscious control: Perceptual and conceptual fluency effects caused by incidental ad exposure. *Journal of Consumer Research*, 26(1), 16–36.
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6(12), 1317–1322.
- Shiv, B., Carmon, Z., & Ariely, D. (2005). Placebo effects of marketing actions: Consumers may get what they pay for. *Journal of Marketing Research*, 42(4), 383–393.
- Shiv, B., Loewenstein, G., & Bechara, A. (2005). Investment behavior and the negative side of emotion. *Psychological Science*, 16(6), 435–439.
- Simonson, I. (2005). In defense of consciousness: The role of conscious and unconscious inputs in consumer choice. *Journal of Consumer Psychology*, 15(3), 211–217.

- Small, D. M., Gregory, M. D., Mak, Y. E., Gitelman, D., Mesulam, M. M., & Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*, 39(4), 701–711.
- Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C., & Jones-Gotman, M. (2001). Changes in brain activity related to eating chocolate: From pleasure to aversion. *Brain*, 124(9), 1720–1733.
- Squire, L. R., & Zola, S. M. (1996). Memory, memory impairment, and the medial temporal lobe. *Cold Spring Harbor Symposia on Quantitative Biology*, 61, 185–195.
- Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences (USA)*, 93(24), 13515–13522.
- Squire, L. R., & Zola, S. M. (1998). Episodic memory, semantic memory, and amnesia. *Hippocampus*, 8(3), 205–211.
- Stewart, A. J., Pickering, M. J., & Sturt, P. (2004). Using eye movements during reading as an implicit measure of acceptability of brand extensions. *Applied Cognitive Psychology*, 18, 697–709.
- Swaminathan, V., Page, K. L., & Gurhan-Canli, Z. (2007). "My" brand or "our" brand: The effects of brand relationship dimensions and self-construal on brand evaluations. *Journal of Consumer Research*, 34, 248–259.
- Synodinos, N. E. (1988). Review and appraisal of subliminal perception within the context of signal detection theory. *Psychology and Marketing*, 5(4), 317–336.
- Talmy, D., Dayan, P., Kiebel, S. J., Frith, C. D., & Dolan, R. J. (2009). How humans integrate the prospects of pain and reward during choice. *Journal of Neuroscience*, 29(46), 14617–14626.
- Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14), 1–17.
- Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, 135(2), 77–99.
- Theus, K. T. (1994). Subliminal advertising and the psychology of processing unconscious stimuli: A review of research. *Psychology and Marketing*, 11(3), 271–290.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12, 97–136.
- Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences of the United States of America*, 91(6), 2016–2020.
- Tusche, A., Bode, S., & Haynes, J. D. (2010). Neural responses to unattended products predict later consumer choices. *Journal of Neuroscience*, 30(23), 8024–8031.
- Van Den Bergh, B., Dewitte, S., & Warlop, L. U. K. (2008). Bikinis instigate generalized impatience in intertemporal choice. *Journal of Consumer Research*, 35(1), 85–97.
- Van der Lans, R., Pieters, R., & Wedel, M. (2008). Competitive brand salience. *Marketing Science*, 27(5), 922–931.
- Van Osselaer, S. M., & Janiszewski, C. (2001). Two ways of learning brand associations. *Journal of Consumer Research*, 28, 202–223.
- van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance*, 30(4), 746–759.
- Vranić, A., & Hromatko, I. (2008). Content-specific activational effects of estrogen on working memory performance. *The Journal of General Psychology*, 135(3), 323–336.
- Waber, R. L., Shiv, B., Carmon, Z., & Ariely, D. (2008). Commercial features of placebo and therapeutic efficacy. *JAMA*, 299(9), 1016–1017.
- Wadhwa, M., Shiv, B., & Nowlis, S. M. (2008). A bite to whet the reward appetite: The influence of sampling on reward-seeking behaviors. *Journal of Marketing Research*, 45(4), 403–413.
- Wallis, J. D., & Miller, E. K. (2003). Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *European Journal of Neuroscience*, 18(7), 2069–2081.
- Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Belknap Press of Harvard University Press.
- Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break? *Psychological Science*, 14(5), 520–524.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6), 495–501.
- Xianchi, D., Brendl, C. M., & Ariely, D. (2010). Wanting, liking, and preference construction. *Emotion*, 10(3), 324–334.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.
- Yoon, C., Gutchess, A. H., Feinberg, F., & Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *Journal of Consumer Research*, 33(1), 31–40.
- Zajonc, R. B., & Markus, H. (1985). Must all affect be mediated by cognition? *Journal of Consumer Research*, 12(3), 363–364.
- Zaltman, G. (2000). Consumer researchers: Take a hike! *Journal of Consumer Research*, 26(4), 423–428.
- Zhu, X., Wang, X., Parkinson, C., Cai, C., Gao, S., & Hu, P. (2010). Brain activation evoked by erotic films varies with different menstrual phases: An fMRI study. *Behavioural Brain Research*, 206(2), 279–285.
- Zink, C. F., Pagnoni, G., Chappelow, J., Martin-Skurski, M., & Berns, G. S. (2006). Human striatal activation reflects degree of stimulus saliency. *Neuroimage*, 29(3), 977–983.
- Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., & Berns, G. S. (2003). Human striatal response to salient nonrewarding stimuli. *Journal of Neuroscience*, 23(22), 8092–8097.
- Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron*, 42(3), 509–517.

A framework for studying the neurobiology of value-based decision making

Antonio Rangel*, Colin Camerer* and P. Read Montague[†]

Abstract | Neuroeconomics is the study of the neurobiological and computational basis of value-based decision making. Its goal is to provide a biologically based account of human behaviour that can be applied in both the natural and the social sciences. This Review proposes a framework to investigate different aspects of the neurobiology of decision making. The framework allows us to bring together recent findings in the field, highlight some of the most important outstanding problems, define a common lexicon that bridges the different disciplines that inform neuroeconomics, and point the way to future applications.

Value-based decision making is pervasive in nature. It occurs whenever an animal makes a choice from several alternatives on the basis of a subjective value that it places on them. Examples include basic animal behaviours, such as bee foraging, and complicated human decisions, such as trading in the stock market. Neuroeconomics is a relatively new discipline that studies the computations that the brain carries out in order to make value-based decisions, as well as the neural implementation of those computations. It seeks to build a biologically sound theory of how humans make decisions that can be applied in both the natural and the social sciences.

The field brings together models, tools and techniques from several disciplines. Economics provides a rich class of choice paradigms, formal models of the subjective variables that the brain needs to compute to make decisions, and some experimental protocols for how to measure these variables. Psychology provides a wealth of behavioural data that shows how animals learn and choose under different conditions, as well as theories about the nature of those processes. Neuroscience provides the knowledge of the brain and the tools to study the neural events that attend decision making. Finally, computer science provides computational models of machine learning and decision making. Ultimately, it is the computations that are central to uniting these disparate levels of description, as computational models identify the kinds of signals and signal dynamics that are required by different value-dependent learning and decision problems. However, a full understanding of choice will require a description at all these levels.

In this Review we propose a framework for thinking about decision making. It has three components: first, it divides decision-making computations into five types; second, it shows that there are multiple types of valuation systems; and third, it incorporates modulating variables that affect the different valuation processes. This framework will allow us to bring together recent findings in the field, highlight some of the most important outstanding problems, define a common lexicon that bridges the different disciplines that inform neuroeconomics, and point the way to future applications. The development of a common lexicon is important because a lot of confusion has been introduced into the literature on the neurobiology of decision making by the use of the unqualified terms 'reward' and 'value'; as shown in the Review, these terms could apply to very different computations.

Computations involved in decision making

The first part of the framework divides the computations that are required for value-based decision making into five basic processes (FIG. 1). The categorization that we propose is based on existing theoretical models of decision making in economics, psychology and computer science^{1–3}. Most models in these disciplines assume, sometimes implicitly, that all of these processes are carried out every time an animal makes a value-based decision.

The first process in decision making involves the computation of a representation of the decision problem. This entails identifying internal states (for example, hunger level), external states (for example, threat level)

*Division of the Humanities and Social Sciences (HSS) and Computational and Neural Systems Program, California Institute of Technology, Pasadena, California 91125, USA. [†]Department of Neuroscience, Computational Psychiatry Unit, Baylor College of Medicine, Houston, Texas 77030, USA.
Correspondence to A.R.
e-mail:
rangel@hss.caltech.edu
doi:10.1038/nrn2357
Published online 11 June 2008

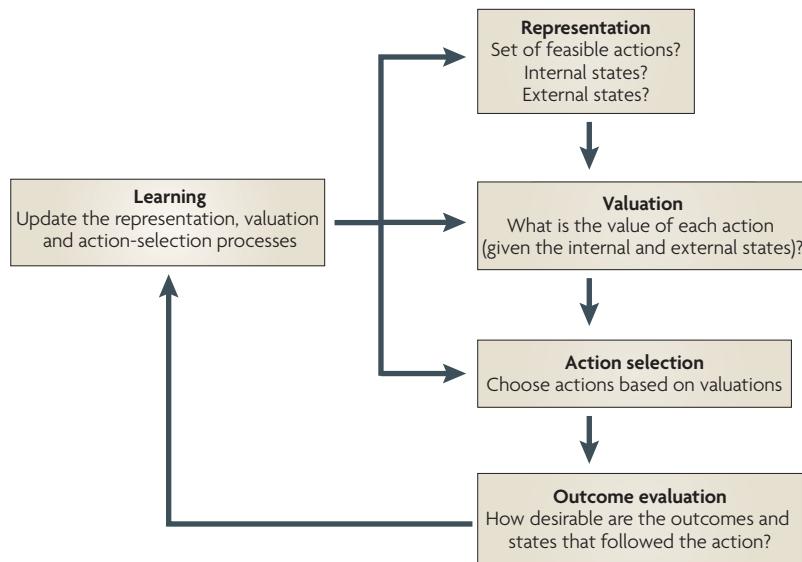


Figure 1 | Basic computations involved in making a choice. Value-based decision making can be broken down into five basic processes: first, the construction of a representation of the decision problem, which entails identifying internal and external states as well as potential courses of action; second, the valuation of the different actions under consideration; third, the selection of one of the actions on the basis of their valuations; fourth, after implementing the decision the brain needs to measure the desirability of the outcomes that follow; and finally, the outcome evaluation is used to update the other processes to improve the quality of future decisions.

and potential courses of action (for example, pursue prey). In the second process, the different actions that are under consideration need to be assigned a value (valuation). In order to make appropriate decisions, these values have to be reliable predictors of the benefits that are likely to result from each action. Third, the different values need to be compared for the animal to be able to make a choice (action selection). Fourth, after implementing the decision, the brain needs to measure the desirability of the outcomes. Finally, these feedback measures are used to update the other processes to improve the quality of future decisions (learning).

The five categories are not rigid, and many questions remain about how well they match the computations that are made by the brain. For example, it is not known whether valuation (step 2 in our model) must occur before action selection (step 3), or whether both computations are performed in parallel. Nevertheless, the taxonomy is conceptually useful because it breaks down the decision-making process into testable constituent processes, it organizes the neuroeconomics literature in terms of the computations that are being studied, and it makes predictions about the neurobiology of decision making, such as the hypothesis that the brain must encode distinct value signals at the decision and outcome stages, and the hypothesis that the brain computes a value signal for every course of action under consideration.

Representation

The representation process plays an essential part in decision making by identifying the potential courses of action that need to be evaluated, as well as the internal

and external states that inform those valuations. For example, the valuation that a predator assigns to the action ‘chasing prey’ is likely to depend on its level of hunger (an internal state) as well as the conditions of the terrain (an external variable). Unfortunately, little is known about the computational and neurobiological basis of this step. Basic open questions include: how does the brain determine which actions to assign values to, and thus consider in the decision-making process, and which actions to ignore? Is there a limit to the number of actions that animals can consider at a time? How are internal and external states computed? How are the states passed to the valuation mechanisms described below?

Valuation at the time of choice

On the basis of a sizable body of animal and human behavioural evidence, several groups have proposed the existence of three different types of valuation systems: Pavlovian, habitual and goal-directed systems^{4–6} (BOX 1). These systems are sometimes in agreement but often in conflict (see section on action selection). It is important to emphasize that the precise neural basis of these three distinct valuation systems is yet to fully be established. Although the evidence described below points to neural dissociations between some of the components of the three hypothetical systems, it is possible that they do not map directly onto completely separate neural systems^{6–9}. In fact, it is likely that they share common elements. Moreover, even the exact nature and number of valuation systems is still being debated. Nevertheless, conceptually the three systems provide a useful operational division of the valuation problem according to the style of the computations that are performed by each.

Pavlovian systems. Pavlovian systems assign values to a small set of behaviours that are evolutionarily appropriate responses to particular environmental stimuli. Typical examples include preparatory behaviours (such as approaching cues that predict the delivery of food) and consummatory responses to a reward (such as pecking at a food magazine). Analogously, cues that predict a punishment or the presence of an aversive stimulus can lead to avoidance behaviours. We refer to these types of behaviours as Pavlovian behaviours, and to the systems that assign value to them as the Pavlovian valuation systems.

Many Pavlovian behaviours are innate, or ‘hard-wired’, responses to specific predetermined stimuli. However, with sufficient training animals can also learn to deploy them in response to other stimuli. For example, rats and pigeons learn to approach lights that predict the delivery of food. An important difference between Pavlovian systems and the other two systems is that Pavlovian systems assign value to only a small set of ‘prepared’ behaviours and thus have a limited behavioural repertoire. Nonetheless, a wide range of human behaviours that have important economic consequences might be controlled by Pavlovian systems, such as overeating in the presence of food, behaviours displayed in people with obsessive-compulsive disorders (OCDs) and, perhaps, harvesting immediate smaller rewards at the expense of delayed larger rewards^{5,9}.

Box 1 | Examples of behaviours driven by different valuation systems

Valuation system	Valence	
	Appetitive (rewards)	Avoidance (punishments)
Pavlovian	Eat all food on plate	Cross street upon seeing dangerous person
	Reward obtained: food	Punishment avoided: potential harm
Habitual	Morning cup of coffee	Drive usual route to work
	Reward obtained: stimulant	Punishment avoided: traffic
Goal-directed	Movie selection	Go for a run
	Reward obtained: entertainment	Punishment avoided: obesity

Behaviour can be driven by different valuation systems. These systems can operate in the domain of rewards (that is, appetitive outcomes) and punishments (that is, aversive outcomes). Although the exact number of valuation systems and their scope remain to be determined, it is known that behaviour can be influenced by Pavlovian, habitual and goal-directed evaluators. The table contains examples of behaviours that are characteristic of each system. Consummatory actions, such as eating food that is within reach, are assigned a high value by the Pavlovian system regardless of the state of hunger. Routine actions, such as having a cup of coffee in the morning, are assigned a high value by the habitual system regardless of that morning's particular needs. Choices that are made infrequently, such as which movie to see, are assigned values by the goal-directed system.

At first glance, Pavlovian behaviours look like automatic, stimulus-triggered responses, and not like instances of value-based choice. However, as Pavlovian responses can be interrupted by other brain systems, they must be assigned something akin to a 'value' so that they can compete with the actions that are favoured by the other valuation systems.

Characterizing the computational and neural basis of Pavlovian systems has so far proven difficult. This is due in part to the fact that there might be multiple Pavlovian controllers, some of which might be responsible for triggering outcome-specific responses (for example, pecking at food or licking at water) whereas others might be responsible for triggering more general valence-dependent responses (for example, approaching for positive outcomes or withdrawing from negative ones).

The neural bases of active and passive Pavlovian responses to negative stimuli seem to have specific and spatial organizations along an axis of the dorsal periaqueductal grey¹⁰. With respect to valence-dependent responses, studies that used various species and methods suggested that a network that includes the basolateral amygdala, the ventral striatum and the orbitofrontal cortex (OFC) underlies the learning processes through which neutral stimuli become predictive of the value of outcomes^{11,12}. In particular, the amygdala has been shown to play a crucial part in influencing some Pavlovian responses^{8,13–15}. Specifically, the central nucleus of the amygdala, through its connections to the brainstem nuclei and the core of the nucleus accumbens, seems to be involved in nonspecific preparatory responses, whereas the basolateral complex of the amygdala seems to be involved in more specific responses through its connections to the hypothalamus and the periaqueductal grey.

Valence

The appetitive or aversive nature of a stimulus.

It is not currently known how many Pavlovian systems exist or how they interact with each other. Other important questions are whether there is a common carrier of Pavlovian value and, if so, how it is encoded; whether learning is possible within these systems; and how Pavlovian systems interact with the other valuation systems — for example, in phenomena such as Pavlovian-instrumental transfer⁴.

Habit systems. In contrast to Pavlovian systems, which value only a small set of responses, habit systems can learn, through repeated training, to assign values to a large number of actions. Habit-valuation systems have a number of key characteristics. First, they learn to assign values to stimulus-response associations (which indicate the action that should be taken in a particular state of the world), on the basis of previous experience, through a process of trial-and-error (see the learning section below). Second, subject to some technical qualifications, habit systems learn to assign a value to actions that is commensurate with the expected reward that these actions generate, as long as sufficient practice is provided and the environment is sufficiently stable^{3,6,16}. Third, because values are learned by trial-and-error, habit systems are believed to learn relatively slowly. As a consequence, they might forecast the value of actions incorrectly immediately after a change in the action-reward contingencies. Finally, these systems rely on 'generalization' when assigning action values in novel situations. For example, a rat that has learned to lever-press for liquids in response to a sound cue might respond with a similar behaviour when first exposed to a light cue. We refer to the actions that are controlled by these systems as 'habits' and the values that they compute as 'habit values'. Examples of habits include a smoker's desire to have a cigarette at particular times of day (for example, after a meal) and a rat's tendency to forage in a cue-dependent location after sufficient training.

Studies using several species and methods suggest that the dorsolateral striatum might play a crucial part in the control of habits^{17,18}. As discussed below, the projections of dopamine neurons into this area are believed to be important for learning the value of actions. Furthermore, it has been suggested that stimulus-response representations might be encoded in cortico-thalamic loops¹⁸. Lesion studies in rats have shown that the infralimbic cortex is necessary for the establishment and deployment of habits^{19,20}.

There are many open questions regarding habit systems. Are there multiple habit systems? How do habitual systems value delayed rewards? What are the limits on the complexity of the environments in which the habit system can learn to compute adequate action values? How does the system incorporate risk and uncertainty? How much generalization is there from one state to another in this system (for example, from hunger to thirst)?

Goal-directed systems. In contrast to the habit system, the goal-directed system assigns values to actions by computing action-outcome associations and then evaluating the rewards that are associated with the different outcomes.

Under ideal conditions, the value that is assigned to an action equals the average reward to which it might lead. We refer to values computed by this system as ‘goal values’ and to the actions that it controls as ‘goal-directed behaviours’. An example of a goal-directed behaviour is the decision of what to eat at a new restaurant.

Note that an important difference between habitual and goal-directed systems has to do with how they respond to changes in the environment. Consider, for example, the valuations made by a rat that has learned to press a lever to obtain food, after it is fed to satiation. The goal-directed system has learned to associate the action ‘lever-press’ with the outcome ‘food’ and thus assigns a value to the lever-press that is equal to the current value of food — which in this example is low because the animal has been fed to satiation. By contrast, the habit system assigns a high value to the lever-press because this is the value that it learned during the pre-satiation training. Thus, the goal-directed system updates the value of an action as soon as the value of its outcome changes, whereas the habit system does not.

To carry out the necessary computations, the goal-directed system needs to store action–outcome and outcome–value associations. Unfortunately, relatively little is known about the neural basis of these processes. Several rat lesion studies suggest that the dorsomedial striatum has a role in the learning and expression of action–outcome associations²¹, whereas the OFC might be responsible for encoding outcome–value associations. Consistent with this, monkey electrophysiology studies have found appetitive goal-value signals in the OFC and in the dorsolateral prefrontal cortex (DLPFC)^{22–25}. Electrophysiology experiments in rats point to the same conclusion²⁶. In a further convergence of findings across methods and species, human functional MRI (fMRI) studies have shown that blood-oxygen-level-dependent (BOLD) activity in the medial OFC^{27–31} and the DLPFC²⁸ correlates with behavioural measures of appetitive goal values, and individuals with damage to the medial OFC have problems making consistent appetitive choices³². Several lines of evidence from these various methods also point to an involvement of the basolateral amygdala and the mediodorsal thalamus (which, in combination with the DLPFC, form a network that Balleine has called the “associative cortico-basal-ganglia loop” (REF. 17)).

Several questions regarding this system remain unanswered. Are there specialized goal-directed systems for reward and punishment, and for different types of goals? How are action–outcome associations learned? How does the goal-directed system assign value to familiar and unfamiliar outcomes? How are action–outcome associations activated at the time that a choice has to be made?

For complex economic choices (such as choosing among detailed health-care plans), we speculate that, in humans, propositional logic systems have a role in constructing associations that are subsequently evaluated by the goal-directed system. For example, individuals might use a propositional system to try to forecast the consequences of a particular action, which are then

evaluated by the goal-directed system. This highlights a limitation of the goal-directed system: the quality of its valuations is limited by the quality of the action–outcome associations that it uses.

Outstanding issues. Some general, important questions regarding the different valuation systems remain unanswered. First, are there multiple Pavlovian, habitual and goal-directed valuation systems, with each system specializing in particular classes of actions (in the case of the Pavlovian and habit systems) or outcomes (in the case of the goal-directed system)? For example, consider a dieter who is offered a tasty dessert at a party. If this is a novel situation, it is likely to be evaluated by the goal-directed system. The dieter is likely to experience conflict between going for the taste of the dessert and sticking to his health goals. This might entail a conflict between two goal-directed systems, one that is focused on the evaluation of immediate taste rewards and one that is focused on the evaluation of long-term outcomes. Second, are there more than three valuation systems? Lengyel and Dayan^{5,33} have proposed the existence of an additional, ‘episodic’ system. At this point it is unclear how such a system differs both conceptually and neurally from the goal-directed system. Third, how does the brain implement the valuation computations of the different systems? Finally, how do long-term goals, cultural norms and moral considerations get incorporated into the valuation process? One possibility is that the habit and goal-directed systems treat violations of these goals and cultural and moral rules as aversive outcomes, and that compliance with them is treated as a rewarding outcome³⁴. However, this can be the case only if the brain has developed the capacity to incorporate social and moral considerations into its standard valuation circuitry. Another possibility is that there are separate valuation systems for these types of considerations that are yet to be discovered.

Modulators of the valuation systems

Several factors can affect the values that the Pavlovian, habitual and goal-directed systems assign to actions. For example, the value that is assigned to an action might depend on the riskiness of its associated payoffs, the delay with which those payoffs occur and the social context. We refer to these types of variables as value modulators. Importantly, modulators might have different effects in each of the valuation systems. In this section we focus on the impact of risk and delay on the goal-directed valuation system, as most of the existing evidence pertains to this system. For reviews on social modulators, see REFS 35,36.

Risk and uncertainty. All decisions involve some degree of risk, in the sense that action–outcome associations are probabilistic (BOX 2). We refer to an action that has uncertain rewards as a ‘prospect’. In order to make good decisions, the goal-directed system needs to take into account the likelihood of the different outcomes. Two hotly debated questions are: first, what are the computations that the goal-directed system uses to incorporate

Propositional logic system
A cognitive system that makes predictions about the world on the basis of known pieces of information.

Statistical moments
Properties of a distribution,
such as mean and variance.

risks into its valuations; and second, how does the brain implement such computations³⁷?

Early human neuroimaging studies in this topic identified some of the areas that are involved in making risky decisions, but were not able to characterize the nature of the computations made by these systems^{38–41}. Currently, two main competing views regarding the nature of such computations are being tested. The first view, which is widely used in financial economics and behavioural ecology,

asserts that the brain assigns value to prospects by first computing its statistical moments (such as its expected magnitude, its variance or coefficient of variation, and its skewness) and then aggregating them into a value signal^{42,43}. The second view, which is widely used in other areas of economics and in psychology, asserts that the value is computed using either expected-utility theory (EU) or prospect theory (PT) (BOX 2). In this case the brain needs to compute a utility value for each potential outcome, which is then weighted by a function of the probabilities.

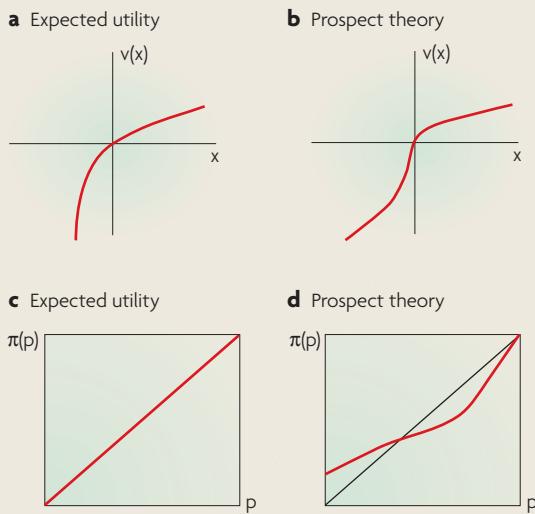
Decisions that result from an EU or PT valuation function can be approximated by a weighted sum of the prospects' statistical moments (and vice versa). This makes it difficult to distinguish the two views on the basis of behavioural data alone. Neuroimaging studies can provide important insights, although the debate between the two views has not yet been settled. In agreement with the first view, a number of recent human fMRI studies have found activity that is consistent with the presence of expected value signals in the striatum^{44,45} and the medial OFC⁴⁶, and activity that is consistent with risk signals (as measured by the mathematical variance of the prospects) in the striatum^{44,47}, the insula^{46,48} and the lateral OFC⁴⁵. Similar risk and expected-value signals have been found in the midbrain dopamine system in electrophysiology studies in non-human primates⁴⁹. Expected-value signals (BOX 2) have also been found in the lateral intraparietal cortex in non-human primate electrophysiology experiments⁵⁰. Consistent with the second view, a recent human fMRI study found evidence for a PT-like value signal in a network that includes the ventral and dorsal striatum, the ventromedial and ventrolateral prefrontal cortex, the anterior cingulate cortex (ACC) and some midbrain dopaminergic regions²⁷. The existence of evidence that is consistent with both views presents an apparent puzzle. A potential resolution that should be explored in future studies is that the striatal-prefrontal network might integrate the statistical moments that are encoded elsewhere into a value signal that exhibits EU- or PT-like properties.

In many circumstances, decision makers have incomplete knowledge of the risk parameters — a situation known as ambiguity that is different from the pure risk case in which the probabilities are known. Human behavioural studies have shown that people generally have an aversion to choices that are ambiguous⁵¹, which suggests that a parameter that measures the amount of ambiguity might be encoded in the brain and might be used to modulate the value signal. Some preliminary human fMRI evidence points to the amygdala, the OFC⁵² and the anterior insula⁵³ as areas where such a parameter might be encoded.

Some issues regarding risk and valuation are still unclear. First, little is known about how risk affects the computation of value in Pavlovian and habitual systems. For example, most reinforcement learning models (see below) assume that the habit learning system encodes a value signal that incorporates expected values but not risks. This assumption, however, has not been thoroughly tested. Second, little is known about how the brain learns the risk parameters. For example, some behavioural

Box 2 | Risk modulators of value in the goal-directed system

Many decisions involve the valuation of rewards and costs that occur probabilistically, often called 'prospects'. There are two dominant theories in economics about how valuation systems incorporate probability in the assignment of value. In expected-utility theory (EU), the value of a prospect equals the sum of the value of the individual outcomes, $v(x)$, weighted by their objective probability, $p(x)$, which is given by $\sum_x p(x)v(x)$. Under some special assumptions on



the function $v(\cdot)$, which are popular in the study of financial markets, the EU formula boils down to a weighted sum of the expected value and the variance of the prospect⁴². The appeal of EU comes from the fact that it is consistent with plausible normative axioms for decision making, from its mathematical tractability and from its success in explaining some aspects of market behaviour. An alternative approach, called prospect theory (PT), states that the value of a prospect equals $\sum_x \pi(p(x))v(x - r)$, where the values of the outcomes now depend on a reference point, r , and are weighted by a nonlinear function, $\pi(\cdot)$, of the objective probabilities^{122,123}. Reference-dependence can create framing effects (analogous to figure-ground switches in vision), in which different values are assigned to the same prospect depending on which reference point is cognitively prominent. The figure illustrates the usual assumptions that are imposed in the value and probability functions by the two theories. As shown in parts a and c, in EU the value function, $v(\cdot)$, is a concave function of outcomes, and the probability function is the identity function. Note that a special case that is often used in the experimental neuroeconomics literature is $v(x) = x$, which makes the EU function reduce to the expected value of the prospect. The properties of PT are illustrated in parts b and d. The value function is usually revealed by choices to be concave for gains but convex for losses. This assumption is justified by the psychologically plausible assumption of diminished marginal sensitivity to both gains and losses starting from the reference point. PT also assumes that $v(x) < -v(-x)$ for $x > 0$, a property called 'loss-aversion', which leads to a kink in the value function. Part d illustrates the version of PT in which small probabilities are overweighted and large probabilities are underweighted. PT has been successful in explaining some behaviour that was inconsistent with EU theory in behavioural experiments with humans¹²³ and monkeys¹²⁴, as well as economic field evidence¹²⁵.

Neuroeconomists make a distinction between prospects that involve risk and those that involve ambiguity. Risk refers to a situation in which all of the probabilities are known. Ambiguity refers to a situation in which some of the probabilities are unknown. The EU and PT models described above apply to valuation under risk, but not under ambiguity. Several models of valuation under ambiguity have been proposed, but none of them has received strong empirical support^{51,126,127}.

Expected-utility theory
A theory that states that the value of a prospect (or of random rewards) equals the sum of the value of the potential outcomes weighted by their probability.

Prospect theory
An alternative to the expected utility theory that also pertains to how to evaluate prospects.

evidence suggests that habit and goal-directed systems learn about probabilities in different ways and that this leads to different probability weighting by the two systems⁵⁴. Finally, more work is required to better characterize the nature of the computations that are made by the amygdala and the insula in decision making under uncertainty. Preliminary insights suggest that the amygdala might have an asymmetric role in the evaluation of gains and losses. For example, humans with amygdala damage made poor decisions if the decisions involved potential gains, but not if they involved losses⁵⁵, and a related study showed that the amygdala

is differentially activated when subjects decide to take risks for large gains and when they decide to accept a sure loss⁵⁶.

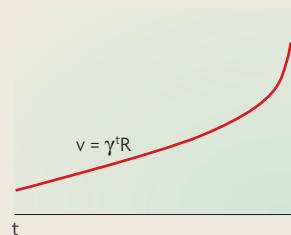
Time discounting. In all real-world situations there is a time lag between decisions and outcomes. From a range of behavioural experiments it is well-established that the goal-directed and habitual systems assign lower values to delayed rewards than to immediate ones; this phenomenon is known as time discounting⁵⁷. The role of time discounting in the Pavlovian system is not as well-understood. As before, we focus on the impact of temporal discounting on the goal-directed system, as this is where most of the studies so far have focused.

The current understanding of time discounting parallels that for risk: two competing views have been proposed and are being tested using a combination of human-behavioural and neuroimaging experiments. One camp interprets the human fMRI evidence using the perspective of dual-process psychological models and has argued that discounting results from the interaction of at least two different neural valuation systems (BOX 3), one with a low discount rate and one with a high discount rate^{58–60}. In this view, the patience that is exhibited by any given individual when making decisions depends on the relative activation of these two systems. In sharp contrast, the other camp has presented human fMRI evidence that suggests that there is a single valuation system that discounts future rewards either exponentially or hyperbolically⁶¹ (BOX 3). As with the situation for risk valuation, this presents an apparent puzzle. A potential reconciliation is that the striatal-prefrontal network might integrate information that is encoded elsewhere in the brain into a single value signal, but that immediate and delayed outcomes might activate different types of information that are used to compute the value. For example, immediate rewards might activate ‘immediacy markers’ that increase the valuation signals in the striatal-prefrontal network. An understanding of these issues is also important from the perspective of brain development. When do value signals get computed in their ‘adult’ form and how do they contribute to choices made by children and adolescents? These and other related questions show that the economic framing of decision making will continue to provide new ways to probe the development and function of choice mechanisms in humans.

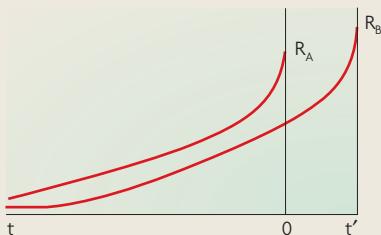
Time discounting remains a fruitful topic of investigation. First, the discounting properties of Pavlovian and habitual systems in humans have not been systematically explored. Second, the inputs to the valuation network are unknown, as is the reason why the aggregation of those inputs produces a hyperbolic-like signal in valuation areas such as the ventral striatum and the medial OFC. Third, the behavioural evidence suggests that discount factors are highly dependent on contextual variables. For example, subjects’ willingness to delay gratification depends on whether the choice is phrased as a “delay” or as a “choice between two points in time” (REF. 62), on how they are instructed to think about the rewards⁶³ and on the subjects’ arousal level⁶⁴. The mechanisms through which such variables affect the valuation process

Box 3 | Temporal modulators of value in the goal-directed system

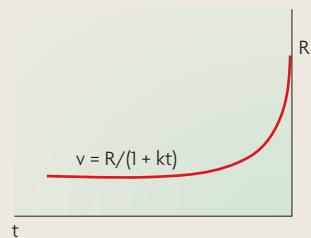
a Exponential



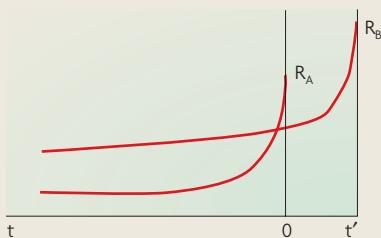
b



c Hyperbolic



d



Many decisions involve the evaluation of rewards and costs that arrive with different delays. Thus, the valuation systems require a mechanism for incorporating the timing of rewards into their computations. Two prominent models of discounting have been proposed in psychology and economics. In the first model, known as hyperbolic discounting, rewards and costs that arrive t units of time in the future are discounted by a factor $1/(1+kt)$. Note that the discount factor is a hyperbolic function of time and that a smaller k is associated with less discounting (that is, more patience). In the second model, known as exponential discounting, the corresponding discount factor is γ^t . Note that a value of γ closer to one is associated with more patience. An important distinction between the two models is illustrated in parts a and c of the figure, which depict the value of a reward of size R t units of time before it arrives. Note that whereas every additional delay is discounted at the same rate (γ) in the exponential case, in hyperbolic discounting initial delays are discounted at a much higher rate and the discount curve flattens out for additional delays.

In most comparative behavioural studies of goal-directed behaviour with adequate statistical power, hyperbolic discount functions always fit the observed behaviour better than exponential functions⁵⁷. Nevertheless, economists and computer scientists find the exponential function appealing because it is the only discount function that satisfies the normative principle of dynamic consistency, which greatly simplifies modelling. This property requires that if a reward, A, is assigned a higher value than another reward, B, at time t , then the same reward is also assigned a higher value when evaluated at any time $t-k$. Under hyperbolic discounting, by contrast, the relative valuation between the two actions depends on when the choice is made. This is known as dynamic inconsistency. Parts b and d of the figure illustrate this difference. They depict the comparative value of a reward, R_A , received at time 0 with a reward, R_B , received at time t' as a function of the time when the rewards are being evaluated. Note that in the exponential case the relative desirability of the two rewards is constant, whereas for the hyperbolic case it depends on the time of evaluation.

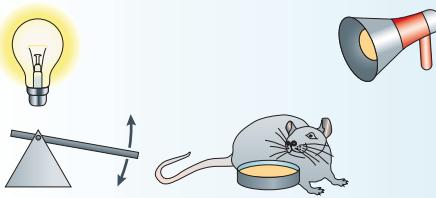
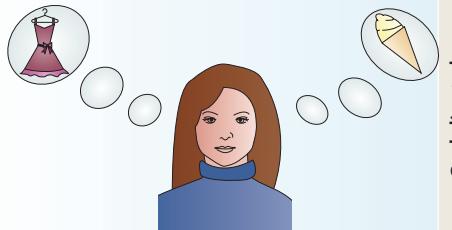
	Pavlovian	Habitual	Goal-directed
Pavlovian	 <p>Example: hungry animal presented with food and electric shock simultaneously Appetitive Pavlovian system: high value for food, low value for escape behaviours Avoidance Pavlovian system: high value for escape behaviours, low value for food</p>	<p>Example: animal rewarded for running away from food Appetitive Pavlovian system: high value for running towards food Avoidance habitual system: high value for running away from food</p>	<p>Example: individual considering taking an extra bite after feeling full Appetitive Pavlovian system: high value for food Health goal-directed system: low value for food</p>
Habitual	 <p>Example: an animal trained to run towards a lever in response to a sound and away from a lever in response to a light being presented with both stimuli Approach habitual system: high value for lever approach Avoidance habitual system: high value for lever avoidance</p>	<p>Example: alcoholic considering having a drink at a bar Appetitive habitual system: high value for drink Avoidance goal-directed system: low value for drink</p>	
Goal-directed	 <p>Example: dieter considering having ice-cream Appetitive goal-directed system: high value for ice-cream Avoidance goal-directed system: low value for ice-cream</p>		

Figure 2 | Conflict between the valuation systems. The different valuation systems are often in agreement. For example, when an individual is hungry at meal time, the Pavlovian, habitual and goal-directed systems assign high value to the consumption of food. However, conflicts between the systems are also common and might lead to poor decision making. This figure provides examples of conflict among the different valuation systems and of conflict among different value signals of the same type.

are unknown. Fourth, several studies have shown that the anticipation of future rewards and punishments can affect subjects' behavioural discount rates^{65,66}. The mechanisms through which anticipation affects valuation are also unknown. Finally, animals make very myopic choices that are consistent with large hyperbolic discount rates^{67–70}. How do humans and animals differ in the way in which they incorporate temporal delays into the valuation process?

Action selection

Even for choices that involve only one of the valuation systems discussed above, options with different values need to be compared in order to make a decision. Little is known about how the brain does this. The only available theoretical models come from the literature on perceptual decision making, which has modelled binary perceptual choices as a race-to-barrier diffusion process^{71–76}. However, it is unclear whether this class of model also applies to value-based decision making and, if so, how the models might be extended to cases of multi-action choice.

Another issue is the competition that arises among the different valuation systems when an animal has to make a choice between several potential actions that are assigned conflicting values (FIG. 2). Some preliminary theoretical

proposals have been made, but the experimental evidence is scarce. Daw *et al.*⁷⁷ have suggested that the brain arbitrates between the habit and goal-directed valuation systems by assigning control to the system that at any given time has the less uncertain estimate of the true value of the actions. As the quality of the estimates that are made by the habit system increases with experience, this means in practice that the habit system should gradually take over from the goal-directed system³⁴. Frank has proposed a neural-network model for choice between appetitive and aversive habitual valuations^{78,79}.

Understanding how the 'control assignment' problem is resolved is important for several reasons. First, as illustrated in FIG. 2 and as emphasized by Dayan *et al.*⁹, many apparently puzzling behaviours are likely to arise as a result of the conflict between the different valuation systems. Second, in most circumstances the quality of decision making depends on the brain's ability to assign control to the valuation system that makes the best value forecasts. For example, it is probably optimal to assign control to the habit system in familiar circumstances, but not in rapidly changing environments. Third, some decision-making pathologies (for example, OCD and overeating) might be due to an inability to assign control to the appropriate system.

Dual-process models

A class of psychological models in which two processes with different properties compete to determine the outcome of a computation.

Race-to-barrier diffusion process

A stochastic process that terminates when the variable of interest reaches a certain threshold value.

There are many important open questions in the domain of action selection. First, in the case of goal-directed decisions, does the brain make decisions by comparing the value of outcomes, of the actions that are necessary to achieve those outcomes, or both? Second, what is the neural basis of the action-selection processes in the Pavlovian, habitual and goal-directed systems? Third, what are the neural mechanisms that are used to arbitrate between the different controllers, and is there a hierarchy of controllers so that some (for example, Pavlovian systems) tend to take precedence over others (for example, goal-directed systems)? Fourth, are there any neural markers that can be reliably used to identify goal-directed or habitual behavioural control?

Outcome evaluation

In order to learn how to make good decisions the brain needs to compute a separate value signal that measures the desirability of the outcomes that were generated by its previous decisions. For example, it is useful for an animal to know whether the last food that it consumed led to illness so that it can know whether it ought to avoid that food in the future.

The computations that are made by the outcome-evaluation system, as well as the neural basis of these computations, are slowly beginning to be understood. The existing evidence comes from several different methods and species. Human fMRI studies have shown that activity in the medial OFC at the time that a reward is being enjoyed correlates with subjective reports about the quality of the experience — this has been shown for olfactory^{80–83}, gustatory^{84–86} and even musical rewards⁸⁷. Moreover, the activity in the medial OFC parallels the reduction in outcome value that one would expect after a subject is fed to satiation^{88,89}. This suggests that the medial OFC might be an area where positive outcome valuations are computed. Interestingly, other human fMRI studies have found positive responses in the medial OFC to the receipt of secondary reinforcers, such as monetary payoffs^{90–92}. Analogous results have been found for negative experiences: in humans, subjective reports of pain intensity correlated with activity in the insula and the ACC^{93,94}.

Animal studies have also provided insight into the neural basis of the outcome-value signal. A recent electrophysiology experiment in monkeys found outcome-value signals in the dorsal ACC⁹⁵. In addition, a series of provocative rat studies showed that it is possible to increase outward manifestations of ‘liking’ in rats (for example, tongue protrusions) by activating the nucleus accumbens and subsets of the ventral pallidum using opioid agonists^{85,96–98}. Interestingly, and consistent with the hypothesis that outcome-evaluation signals play a part in learning, rats that received opioid agonists subsequently consumed more of the reward that was paired with the agonist.

Some recent human fMRI experiments have also provided novel insights into the computational properties of the outcome-value signal. For example, one study showed that activity in the medial OFC in response to an odour depended on whether subjects believed that

they were smelling cheddar cheese or a sweaty sock⁸³. In another study⁹⁹, activity in the medial OFC in response to the consumption of wine depended on beliefs about its price, and a third study⁸⁴ showed that the outcome-evaluation signal after consumption of soda depended on beliefs about its brand. Together, these findings suggest that the outcome-evaluation system is modulated by higher cognitive processes that determine expectancies and beliefs.

Much remains to be understood about the outcome-evaluation system. What network is responsible for computing positive and negative outcome values in different types of domains? How are positive and negative outcome-evaluation signals integrated? How are these signals passed to the learning processes described in the next section? Can they be modulated by variables such as long-term goals, social norms and moral considerations?

Learning

Although some Pavlovian behaviours are innate responses to environmental stimuli, most forms of behaviour involve some form of learning. In fact, in order to make good choices animals need to learn how to deploy the appropriate computations during the different stages of decision making. First, the brain must learn to activate representations of the most advantageous behaviours in every state. This is a non-trivial learning problem given that animals and humans have limited computational power, and yet they can deploy a large number of behavioural responses. Second, the valuation systems must learn to assign to actions values that match their anticipated rewards. Finally, the action-selection processes need to learn how to best allocate control among the different valuation systems.

Of all of these processes, the one that is best-understood is the learning of action values by the habit system. In this area there has been a productive interplay between theoretical models from computer science (BOX 4) and experiments using electrophysiology in rats and monkeys and fMRI in humans. In particular, various reinforcement learning models have been proposed to describe the computations that are made by the habit system¹⁰⁰. The basic idea behind these models is that a prediction-error signal is computed after observing the outcome generated by every choice. The signal is called a prediction error because it measures the quality of the forecast that was implicit in the previous valuation (BOX 4). Every time a learning event occurs, the value of the actions is changed by an amount that is proportional to the prediction error. Over time, and under the appropriate technical conditions, the animal learns to assign the correct value to actions.

The existence of prediction-error-like signals in the brain is one of the best-documented facts in neuroeconomics. Schultz and colleagues initially observed such signals in electrophysiology studies performed in midbrain dopamine neurons of monkeys^{101–106}. The connection between these signals and reinforcement-learning models was made in a series of papers by Montague and colleagues that were published in the 1990s^{103,107}. Since then, several fMRI studies have

Box 4 | Reinforcement learning models action-value learning in the habitual system

Several models from computer science have proved to be useful in modelling how the habitual system learns to assign values to actions. All of these models have the following structure, which is known as a Markovian decision problem: first, the animal can be in a finite set of states and can take a finite set of actions; second, there is a transition function, $T(s,a,s')$, that specifies the probability that state s and action a at one time-step will result in the state s' at the next time-step; and third, at every time-step the animal obtains an action and a state-contingent reward, $r(a,s)$. A behavioural rule in this world (called a policy and denoted by $\pi(s)$) specifies the action that the animal takes in every state. In this world the habitual system needs to solve two problems. First, given a policy, it needs to compute the value of taking every action a in every state s . This is given by

$$Q^\pi(s, a) = E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots | s_t = s, a_t = a, a_{t+1} = \pi(s_{t+1}), \dots]; \quad (1)$$

where r_{t+k} denotes the reward that is received at time $t+k$ and where $\gamma > 0$ is the discount rate. Second, it needs to identify the policy that generates the largest sum of exponentially discounted rewards (see BOX 3) in every state.

How could the habitual system learn $Q^\pi(s, a)$? Let $Q^\pi(s, a)$ denote the estimate that the system has at any point in time. Equation 1 can be rewritten in recursive form as

$$Q^\pi(s, a) = R(s) + \gamma \sum_{s' \in S} T(s, a, s') Q^\pi(s', \pi(s')) \quad (2)$$

Consider an estimator, $\hat{Q}(s, a)$, of $Q^\pi(s, a)$. Note that if $\hat{Q}(s, a)$ does not satisfy this expression, then it is not a good estimate of the value function. Define a prediction error

$$\delta_t = r_t + \gamma \max_{a'} [\hat{Q}(s_{t+1}, a')] - \hat{Q}(s_t, a_t) \quad (3)$$

that is a sample measure of how close the estimate is to satisfying equation 2. If $\delta_t > 0$ the value of the action is overestimated; if $\delta_t < 0$ the value is underestimated. One can then use the prediction error to update the estimates of the action values as follows:

$$\hat{Q}(s_t, a_t) \leftarrow \hat{Q}(s_t, a_t) + \eta \delta_t \quad (4)$$

where η is a number between 0 and 1 that determines the speed of learning. This model is known as Q-learning and it satisfies one important property: subject to some technical conditions, the estimated action values converge to those that are generated by the optimal policy. It then follows that the animal can learn the optimal policy simply by following this algorithm and, at every step of the learning process, selecting the actions with the largest values. Two other variants of this model have been proposed as descriptions of how the habitual system learns. They are known as SARSA and the actor–critic model. They differ from Q-learning on the exact specification of the prediction error and the update rule, but they are based on essentially the same idea. Note that neither SARSA nor the actor–critic model is guaranteed to converge to the optimal policy.

It is worth emphasizing several properties of these learning models. First, they are model-free in the sense that the animal is not assumed to know anything about the transition function or the reward function. Second, they explain a wide range of conditioning behaviours that are associated with the habitual system, such as blocking, overshadowing and inhibitory conditioning. Finally, they are computationally simple in the sense that they do not require the animal to keep track of long sequences of rewards to learn the value of actions.

The reinforcement-learning models described here are often used to describe the process of action-value learning in the habitual system. The algorithms that the Pavlovian and goal-directed systems use to update their values based on feedback from the environment are currently unknown.

shown that, in humans, the BOLD signal in the ventral striatum (an important target of midbrain dopamine neurons) correlates with prediction errors in a wide range of tasks^{29,90,108–113}.

Although the existing evidence suggests that there is a remarkable match between the computational models and the activity of the dopamine system, recent experiments have demonstrated that much remains to be understood. First, a monkey electrophysiology study¹¹⁴ suggested that the phasic firing rates of midbrain dopamine neurons might only encode the positive component of the prediction error (henceforth the ‘positive prediction error’). This raises the question of which brain areas and neurotransmitter systems encode the negative component (henceforth the ‘negative prediction error’), which is also essential for learning. Several possibilities have been proposed. A secondary analysis¹¹⁵ of the monkey

electrophysiology experiment¹¹⁴ suggested that the magnitude of the negative prediction errors might be encoded in the timing of the fire-and-pause patterns of the dopamine cells¹¹⁵. Some human fMRI studies have found a BOLD signal in the amygdala that resembles a negative prediction error¹⁰⁸, but others have failed to replicate this finding and have instead found evidence for both types of prediction error in different parts of the striatum¹¹⁶. In turn, Daw and Dayan¹¹⁷ proposed that the two prediction-error signals are encoded by the phasic responses of two neurotransmitter systems: dopamine for positive prediction errors and serotonin for negative prediction errors. Second, it was shown that midbrain dopamine neurons adjust their firing rates to changes in the magnitude of reward in a way that is inconsistent with the standard interpretation of prediction errors⁴⁹. The exact nature of these adjustments remains an open

Box 5 | From neuroeconomics to computational psychiatry

Sometimes the brain's decision-making processes function so differently from societal norms that we label the ensuing behaviours and perceptions a psychiatric disease. The medical community recognizes and categorizes them according to well-accepted diagnostic criteria that, so far, have relied mostly on collections of behavioural features. Neuroscientists have accumulated a substantial amount of neurobiological data that impinges directly on these illnesses¹²⁸. For example, there are now animal models for nicotine addiction, anxiety, depression and schizophrenia that have produced a veritable flood of data on neurotransmitter systems, receptors and gene expression^{129,130}. Thus, there is a substantial body of biological data and detailed descriptions of the behavioural outcomes, but little is known about what connects them. This situation presents an opportunity for neuroeconomics and other computationally oriented sciences to connect the growing body of biological knowledge to the behavioural end points.

Computational models of reinforcement learning provide a new language for understanding mental illness and a starting point for connecting detailed neural substrates to behavioural outcomes. For example, reinforcement-learning models predict the existence of valuation malfunctions, in which a drug, a disease or a developmental event perturbs the brain's capacity to assign appropriate value to behavioural acts or mental states^{34,131–133}.

Disorders of decision making can also arise at the action-selection stage, especially when there are conflicts among the valuation systems. This presents the possibility of generating a new quantifiable taxonomy of mental-disease states. Interestingly, this set of issues is closely related to the problem of how to think about the 'will' and has applications to addiction, obsessive-compulsive disorder and obesity. These issues relate directly to the idea of executive control and the way that it is affected by mental disease. It is our opinion that future progress in this area will require more computational approaches, because only through such models can competing ideas of executive control be clearly differentiated. Such efforts are already well underway, and various modelling efforts have been applied to executive control and decision making in humans^{79,134,135}.

Another neuroeconomics concept that is ripe for applications to psychiatry is motivation, which is a measure of how hard an animal works in order to retrieve a reward. Disorders of motivation might play an especially important part in mood disorders, such as depression, and in Parkinson's disease^{78,136}.

question⁴³. Finally, a study showed that the habit system can also learn from observing the outcomes of actions that it did not take, as opposed to only being able to learn from direct experience¹¹⁸. This form of 'fictive learning' is not captured by traditional reinforcement-learning models but is common in human strategic learning and suggests that the theory needs to be extended in new directions (to include, among others, imitative learning from observing the actions of others)¹¹⁹.

Other important questions in the domain of value learning include the following: how does the goal-directed system learn the action-outcome and outcome-value representations that it needs to compute action values? What are the limitations of the habit system in situations in which there is a complex credit-assignment problem (because actions and outcomes are not perfectly alternated) and delayed rewards? How does the habit system learn to incorporate internal and external states in its valuations and generalize across them? How do the different learning systems incorporate expected uncertainty about the feedback signals⁴³? To what extent can the different value systems learn by observation as opposed to through direct experience¹²⁰?

The next 5 years and beyond

Although neuroeconomics is a new field and many central questions remain to be answered, rapid progress is being made. As illustrated by the framework provided in this Review, the field now has a coherent lexicon and research aims. The key challenge for neuroeconomics over the next few years is to provide a systematic characterization of the computational and neurobiological basis of the representation, valuation, action-comparison, outcome-valuation and value-learning processes described above. This will prove to be challenging because, as we have seen, at least

three valuation systems seem to be at work, fighting over the control of the decision-making process.

Nevertheless, several welcome developments suggest that the next 5 years will produce significant progress in answering many of the questions outlined here. First, there is the close connection between theory and experiments, and the widespread use of theory-driven experimentation (including behavioural parameters inferred from choices that can be linked across subjects or trials to brain activity). Second, there is the rapid adoption of new technologies, such as fast cyclic voltammetry in freely moving animals¹²¹, which permits quasi-real-time monitoring of neurotransmitter levels for long periods. Third, there is the investigation of decision-making phenomena using different species and experimental methods, which permits more rapid progress.

This is good news, because the range of potential applications is significant. The most important area in which knowledge from neuroeconomics can be applied is psychiatry. Many psychiatric disorders involve a failure of one or more of the decision-making processes described here (BOX 5). A better understanding of these processes should lead to improved diagnosis and treatment. Another area of application is the judicial system. A central question in many legal procedures is how to define and measure whether individuals are in full command of their decision-making faculties. Neuroeconomics has the potential to provide better answers to this question. Similarly, an improved understanding of why people experience failures of self-control should lead to better public-policy interventions in areas ranging from addiction and obesity to savings. The field also has the potential to improve our understanding of how marketing affects decisions and when it should be

Credit-assignment problem
The problem of crediting rewards to particular actions in complex environments.

- regulated. Artificial intelligence is another fertile area of application: a question of particular interest is which features of the brain's decision-making mechanisms are optimal and should be imitated by artificial systems, and which mechanisms can be improved upon. Finally, neuroeconomics might advance our understanding of how to train individuals to become better decision-makers, especially in conditions of extreme time-pressure and large stakes, such as those that arise in policing, in war and in fast-paced financial markets.
1. Busemeyer, J. R. & Johnson, J. G. in *Handbook of Judgment and Decision Making* (eds Koehler, D. & Narvey, N.) 133–154 (Blackwell Publishing Co., New York, 2004).
 2. Mas-Colell, A., Whinston, M. & Green, J. *Microeconomic Theory* (Cambridge Univ. Press, Cambridge, 1995).
 3. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, Massachusetts, 1998).
 4. Dickison, A. & Balleine, B. W. in *Steven's Handbook of Experimental Psychology Vol. 3 Learning, Motivation & Emotion* (ed. Gallistel, C.) 497–533 (Wiley & Sons, New York, 2002).
 5. Dayan, P. in *Better Than Conscious? Implications for Performance and Institutional Analysis* (eds Engel, C. & Singer, W.) 51–70 (MIT Press, Cambridge, Massachusetts, 2008).
 6. Balleine, B. W., Daw, N. & O'Doherty, J. in *Neuroeconomics: Decision-Making and the Brain* (eds Glimcher, P. W., Fehr, E., Camerer, C. & Poldrack, R. A.) 365–385 (Elsevier, New York, 2008).
 7. Bouton, M. E. *Learning and Behavior: A Contemporary Synthesis* (Sinauer Associates, Inc., Sunderland, Massachusetts, 2007). **This book reviews a large amount of evidence pointing to multiple valuation systems being active in value-based decision making.**
 8. Dayan, P. & Seymour, B. in *Neuroeconomics: Decision Making and the Brain* (eds Glimcher, P. W., Camerer, C. F., Fehr, E. & Poldrack, R. A.) 175–191 (Elsevier, New York, 2008).
 9. Dayan, P., Niv, Y., Seymour, B. & Daw, N. D. The misbehavior of value and the discipline of the will. *Neural Netw.* **19**, 1153–1160 (2006). **This paper provided several models of how "pathological behaviours" can arise from the competition process between Pavlovian, habitual and goal-directed valuation systems.**
 10. Keay, K. A. & Bandler, R. Parallel circuits mediating distinct emotional coping reactions to different types of stress. *Neurosci. Biobehav. Rev.* **25**, 669–678 (2001).
 11. Cardinal, R. N., Parkinson, J. A., Hall, J. & Everitt, B. J. Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* **26**, 321–352 (2002).
 12. Holland, P. C. & Gallagher, M. Amygdala-frontal interactions and reward expectancy. *Curr. Opin. Neurobiol.* **14**, 148–155 (2004).
 13. Fendt, M. & Fanselow, M. S. The neuroanatomical and neurochemical basis of conditioned fear. *Neurosci. Biobehav. Rev.* **23**, 743–760 (1999).
 14. Adams, D. B. Brain mechanisms of aggressive behavior: an updated review. *Neurosci. Biobehav. Rev.* **30**, 304–318 (2006).
 15. Niv, Y. in *Neuroscience* (Hebrew University, Jerusalem, 2007).
 16. Dayan, P. & Abbott, L. R. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, Massachusetts, 1999).
 17. Balleine, B. W. Neural bases of food-seeking: affect, arousal and reward in corticostriatal limbic circuits. *Physiol. Behav.* **86**, 717–730 (2005). **This important paper reviews a large amount of evidence pointing to multiple valuation systems being active in value-based decision making.**
 18. Yin, H. H. & Knowlton, B. J. The role of the basal ganglia in habit formation. *Nature Rev. Neurosci.* **7**, 464–476 (2006).
 19. Killcross, S. & Coutureau, E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex* **13**, 400–408 (2003).
 20. Coutureau, E. & Killcross, S. Inactivation of the infralimbic prefrontal cortex reinstates goal-directed responding in overtrained rats. *Behav. Brain Res.* **146**, 167–174 (2003).
 21. Yin, H. H., Knowlton, B. J. & Balleine, B. W. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* **22**, 505–512 (2005).
 22. Wallis, J. D. & Miller, E. K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**, 2069–2081 (2003).
 23. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006). **This paper showed that neurons in the monkey OFC encode the goal value of individual rewarding objects (for example, different liquids) irrespective of the action that needs to be taken to obtain them.**
 24. Wallis, J. D. Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci.* **30**, 31–56 (2007).
 25. Barracough, D. J., Conroy, M. L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neurosci.* **7**, 404–410 (2004).
 26. Schoenbaum, G. & Roesch, M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron* **47**, 633–636 (2005).
 27. Tom, S. M., Fox, C. R., Trepel, C. & Poldrack, R. A. The neural basis of loss aversion in decision-making under risk. *Science* **315**, 515–518 (2007). **This fMRI study showed that the striatal-OFC network encodes a value signal at the time of the goal-directed choice that is consistent with the properties of PT. Furthermore, the study presented evidence that suggests that both the appetitive and the aversive aspects of goal-directed decisions might be encoded in a common valuation network.**
 28. Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* **27**, 9984–9988 (2007).
 29. Hare, T., O'Doherty, J., Camerer, C. F., Schultz, W. & Rangel, A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* (in press).
 30. Paulus, M. P. & Frank, L. R. Ventromedial prefrontal cortex activation is critical for preference judgments. *Neuroreport* **14**, 1311–1315 (2003).
 31. Erk, S., Spitzer, M., Wunderlich, A. P., Galley, L. & Walter, H. Cultural contexts modulate reward circuitry. *Neuroreport* **13**, 2499–2503 (2002).
 32. Fellows, L. K. & Farah, M. J. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb. Cortex* **17**, 2669–2674 (2007).
 33. Lengyel, M. & Dayan, P. Hippocampal contributions to control: the third way. *NIPS* [online] http://books.nips.cc/papers/files/nips20/NIPS2007_0927.pdf (2007).
 34. Montague, P. R. *Why Choose This Book?* (Dutton, 2006).
 35. Fehr, E. & Camerer, C. F. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* **11**, 419–427 (2007).
 36. Lee, D. Game theory and neural basis of social decision making. *Nature Neurosci.* **11**, 404–409 (2008).
 37. Platt, M. L. & Huettel, S. A. Risky business: the neuroeconomics of decision making under uncertainty. *Nature Neurosci.* **11**, 398–403 (2008).
 38. Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S. & Stein, M. B. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* **19**, 1439–1448 (2003).
 39. Leland, D. S. & Paulus, M. P. Increased risk-taking decision-making but not altered response to punishment in stimulant-using young adults. *Drug Alcohol Depend.* **78**, 83–90 (2005).
 40. Paulus, M. P. et al. Prefrontal, parietal, and temporal cortex networks underlie decision-making in the presence of uncertainty. *Neuroimage* **13**, 91–100 (2001).
 41. Huettel, S. A., Song, A. W. & McCarthy, G. Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *J. Neurosci.* **25**, 3304–3311 (2005).
 42. Bossaerts, P. & Hsu, M. in *Neuroeconomics: Decision Making and the Brain* (eds Glimcher, P. W., Camerer, C. F., Fehr, E. & Poldrack, R. A.) 351–364 (Elsevier, New York, 2008).
 43. Preuschoff, K. & Bossaerts, P. Adding prediction risk to the theory of reward learning. *Ann. NY Acad. Sci.* **1104**, 135–146 (2007).
 44. Preuschoff, K., Bossaerts, P. & Quartz, S. R. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* **51**, 381–390 (2006).
 45. Tobler, P. N., O'Doherty, J. P., Dolan, R. J. & Schultz, W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* **97**, 1621–1632 (2007).
 46. Rolls, E. T., McCabe, C. & Redoute, J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb. Cortex* **18**, 652–663 (2007).
 47. Dreher, J. C., Kohn, P. & Berman, K. F. Neural coding of distinct statistical properties of reward information in humans. *Cereb. Cortex* **16**, 561–573 (2006).
 48. Preuschoff, K., Quartz, S. R. & Bossaerts, P. Human insula activation reflects prediction errors as well as risk. *J. Neurosci.* **28**, 2745–2752 (2008). **This fMRI study shows that the human insula encodes risk-prediction errors that could be used to learn the riskiness of different options and that are complementary to reward-prediction errors.**
 49. Tobler, P. N., Fiorillo, C. D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
 50. Platt, M. L. & Glimcher, P. W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
 51. Camerer, C. F. & Weber, M. Recent developments in modelling preferences: uncertainty and ambiguity. *J. Risk Uncertain.* **5**, 325–370 (1992).
 52. Hsu, M., Bhatt, M., Adolphs, R., Tranel, D. & Camerer, C. F. Neural systems responding to degrees of uncertainty in human decision-making. *Science* **310**, 1680–1683 (2005).
 53. Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T. & Platt, M. L. Neural signatures of economic preferences for risk and ambiguity. *Neuron* **49**, 765–775 (2006).
 54. Hertwig, R., Barron, G., Weber, E. U. & Erev, I. Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* **15**, 534–539 (2004).
 55. Weller, J. A., Levin, I. P., Shiv, B. & Bechara, A. Neural correlates of adaptive decision making for risky gains and losses. *Psychol. Sci.* **18**, 958–964 (2007).
 56. De Martino, B., Kumaran, D., Seymour, B. & Dolan, R. J. Frames, biases, and rational decision-making in the human brain. *Science* **313**, 684–687 (2006).
 57. Frederick, S., Loewenstein, G. & O'Donoghue, T. Time discounting and time preference: a critical review. *J. Econ. Lit.* **40**, 351–401 (2002).
 58. McClure, S. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. Separate neural systems value immediate and delayed monetary rewards. *Science* **306**, 503–507 (2004). **This fMRI study argued that competing goal-directed valuation systems play a part in decisions that involve choosing between immediate small monetary payoffs and larger but delayed payoffs.**
 59. McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G. & Cohen, J. D. Time discounting for primary rewards. *J. Neurosci.* **27**, 5796–5804 (2007).
 60. Berns, G. S., Laibson, D. & Loewenstein, G. Intertemporal choice - toward an integrative framework. *Trends Cogn. Sci.* **11**, 482–488 (2007).
 61. Kable, J. W. & Glimcher, P. W. The neural correlates of subjective value during intertemporal choice. *Nature Neurosci.* **10**, 1625–1633 (2007). **This fMRI study argued that a single goal-directed valuation system plays a part in decisions that involve choosing between immediate small monetary payoffs and larger but delayed payoffs.**

62. Read, D., Frederick, S., Orsel, B. & Rahman, J. Four score and seven years ago from now: the "date/delay" effect in temporal discounting. *Manage. Sci.* **51**, 1326–1335 (1997).
63. Mischel, W. & Underwood, B. Instrumental ideation in delay of gratification. *Child Dev.* **45**, 1083–1088 (1974).
64. Wilson, M. & Daly, M. Do pretty women inspire men to discount the future? *Proc. Biol. Sci.* **271** (Suppl 4), S177–S179 (2004).
65. Berns, G. S. *et al.* Neurobiological substrates of dread. *Science* **312**, 754–758 (2006).
66. Loewenstein, G. Anticipation and the valuation of delayed consumption. *Econ. J.* **97**, 666–684 (1987).
67. Stevens, J. R., Hallinan, E. V. & Hauser, M. D. The ecology and evolution of patience in two New World monkeys. *Biol. Lett.* **1**, 223–226 (2005).
68. Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* **4**, 267–272 (1961).
69. Mazur, J. E. Estimation of indifference points with an adjusting-delay procedure. *J. Exp. Anal. Behav.* **49**, 37–47 (1988).
70. Corrado, G. S., Sugrue, L. P., Seung, H. S. & Newsome, W. T. Linear/nonlinear-poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* **84**, 581–617 (2005).
71. Newsome, W. T., Britten, K. H. & Movshon, J. A. Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989).
72. Kim, J. N. & Shadlen, M. N. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neurosci.* **2**, 176–185 (1999).
73. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
74. Gold, J. I. & Shadlen, M. N. Banburisms and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* **36**, 299–308 (2002).
75. Gold, J. I. & Shadlen, M. N. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16 (2001).
76. Heekeren, H. R., Marrett, S. & Ungerleider, L. G. The neural systems that mediate human perceptual decision making. *Nature Rev. Neurosci.* **9**, 467–479 (2008).
77. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neurosci.* **8**, 1704–1711 (2005). **This paper proposed a theoretical model of how the brain might assign control to the different goal and habitual systems.**
78. Frank, M. J., Seeberger, L. C. & O'Reilly, R. C. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* **306**, 1940–1943 (2004).
79. Frank, M. J. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural Netw.* **19**, 1120–1136 (2006).
80. de Araujo, I. E., Rolls, E. T., Krriegelbach, M. L., McGlone, F. & Phillips, N. Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *Eur. J. Neurosci.* **18**, 2059–2068 (2003).
81. de Araujo, I. E., Krriegelbach, M. L., Rolls, E. T. & McGlone, F. Human cortical responses to water in the mouth, and the effects of thirst. *J. Neurophysiol.* **90**, 1865–1876 (2003).
82. Anderson, A. K. *et al.* Dissociated neural representations of intensity and valence in human olfaction. *Nature Neurosci.* **6**, 196–202 (2003).
83. de Araujo, I. E., Rolls, E. T., Velazco, M. I., Margot, C. & Cayeux, I. Cognitive modulation of olfactory processing. *Neuron* **46**, 671–679 (2005).
84. McClure, S. M. *et al.* Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* **44**, 379–387 (2004).
85. Krriegelbach, M. L., O'Doherty, J., Rolls, E. T. & Andrews, C. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb. Cortex* **13**, 1064–1071 (2003).
86. Small, D. M. *et al.* Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* **39**, 701–711 (2003).
87. Blood, A. J. & Zatorre, R. J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl Acad. Sci. USA* **98**, 11818–11823 (2001).
88. O'Doherty, J. *et al.* Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport* **11**, 399–403 (2000).
89. Small, D. M., Zatorre, R. J., Dagher, A., Evans, A. C. & Jones-Gotman, M. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain* **124**, 1720–1733 (2001).
90. Breiter, H. C., Aharon, I., Kahneman, D., Dale, A. & Shizgal, P. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* **30**, 619–639 (2001).
91. Knutson, B., Fong, G. W., Adams, C. M., Varner, J. L. & Hommer, D. Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* **12**, 3683–3687 (2001).
92. Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C. & Berns, G. S. Human striatal responses to monetary reward depend on saliency. *Neuron* **42**, 509–517 (2004).
93. Peyron, R. *et al.* Haemodynamic brain responses to acute pain in humans: sensory and attentional networks. *Brain* **122**, 1765–1780 (1999).
94. Davis, K. D., Taylor, S. J., Crawley, A. P., Wood, M. L. & Mikulis, D. J. Functional MRI of pain- and attention-related activations in the human cingulate cortex. *J. Neurophysiol.* **77**, 3370–3380 (1997).
95. Seo, H. & Lee, D. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* **27**, 8366–8377 (2007).
96. Pecina, S., Smith, K. S. & Berridge, K. C. Hedonic hot spots in the brain. *Neuroscientist* **12**, 500–511 (2006).
97. Berridge, K. C. & Robinson, T. E. Parsing reward. *Trends Neurosci.* **26**, 507–513 (2003).
98. Berridge, K. C. Pleasures of the brain. *Brain Cogn.* **52**, 106–128 (2003).
99. Plassmann, H., O'Doherty, J., Shiv, B. & Rangel, A. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl Acad. Sci. USA* **105**, 1050–1054 (2008). **This paper showed that the level of "experienced pleasantness" encoded in the medial OFC at the time of consuming a wine is modulated by subjects' beliefs about the price of the wine that they are drinking.**
100. Montague, P. R., King-Casas, B. & Cohen, J. D. Imaging valuation models in human choice. *Annu. Rev. Neurosci.* **29**, 417–448 (2006).
101. Tremblay, L., Hollerman, J. R. & Schultz, W. Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J. Neurophysiol.* **80**, 964–977 (1998).
102. Hollerman, J. R. & Schultz, W. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neurosci.* **1**, 304–309 (1998).
103. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997). **This paper proposed the connection between the prediction-error component of reinforcement-learning models and the behaviour of dopamine cells.**
104. Mirenowicz, J. & Schultz, W. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* **72**, 1024–1027 (1994).
105. Schultz, W. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2007).
106. Schultz, W. Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr. Opin. Neurobiol.* **14**, 139–147 (2004).
107. Montague, P. R., Dayan, P. & Sejnowski, T. J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
108. Yacubian, J. *et al.* Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* **26**, 9530–9537 (2006).
109. Tanaka, S. C. *et al.* Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neurosci.* **7**, 887–893 (2004).
110. Pagnoni, G., Zink, C. F., Montague, P. R. & Berns, G. S. Activity in human ventral striatum locked to errors of reward prediction. *Nature Neurosci.* **5**, 97–98 (2002).
111. O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H. & Dolan, R. J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329–337 (2003).
112. Knutson, B., Westdorp, A., Kaiser, E. & Hommer, D. fMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage* **12**, 20–27 (2000).
113. Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C. & Fiez, J. A. Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* **84**, 3072–3077 (2000).
114. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
115. Bayer, H. M., Lau, B. & Glimcher, P. W. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* **98**, 1428–1439 (2007).
116. Seymour, B., Daw, N., Dayan, P., Singer, T. & Dolan, R. Differential encoding of losses and gains in the human striatum. *J. Neurosci.* **27**, 4826–4831 (2007).
117. Daw, N. D., Kakade, S. & Dayan, P. Opponent interactions between serotonin and dopamine. *Neural Netw.* **15**, 603–616 (2002).
118. Lohrenz, T., McCabe, K., Camerer, C. F. & Montague, P. R. Neural signature of fictive learning signals in a sequential investment task. *Proc. Natl Acad. Sci. USA* **104**, 9493–9498 (2007).
119. Camerer, C. F. & Chong, J. K. Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* **133**, 177–198 (2007).
120. Olsson, A. & Phelps, E. A. Social learning of fear. *Nature Neurosci.* **10**, 1095–1102 (2007).
121. Montague, P. R. *et al.* Dynamic gain control of dopamine delivery in freely moving animals. *J. Neurosci.* **24**, 1754–1759 (2004).
122. Tversky, A. & Kahneman, D. Advances in prospect theory cumulative representation of uncertainty. *J. Risk Uncertain.* **5**, 297–323 (1992).
123. Kahneman, D. & Tversky, A. Prospect Theory: an analysis of decision under risk. *Econometrica* **4**, 263–291 (1979). **This seminal paper proposed the PT model for goal-directed valuation in the presence of risk and provided some supporting evidence. It is one of the most cited papers in economics.**
124. Chen, K., Lakshminarayanan, V. & Santos, L. How basic are behavioral biases? Evidence from capuchin-monkey trading behavior. *J. Polit. Econ.* **114**, 517–537 (2006).
125. Camerer, C. F. in *Choice, Values, and Frames* (eds Kahneman, D. & Tversky, A.) (Cambridge Univ. Press, Cambridge, 2000).
126. Gilboa, I. & Schmeidler, D. Maximin expected utility with non-unique prior. *J. Math. Econ.* **28**, 141–153 (1989).
127. Ghirardato, P., Maccheroni, F. & Marinacci, M. Differentiating ambiguity and ambiguity attitude. *J. Econ. Theory* **118**, 133–173 (2004).
128. Nestler, E. J. & Charney, D. S. *The Neurobiology of Mental Illness* (Oxford Univ. Press, Oxford, 2004).
129. Kauer, J. A. & Malenka, R. C. Synaptic plasticity and addiction. *Nature Rev. Neurosci.* **8**, 844–858 (2007).
130. Hyman, S. E., Malenka, R. C. & Nestler, E. J. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu. Rev. Neurosci.* **29**, 565–598 (2006).
131. Redish, A. D. & Johnson, A. A computational model of craving and obsession. *Ann. NY Acad. Sci.* **1104**, 324–339 (2007).
132. Redish, A. D. Addiction as a computational process gone awry. *Science* **306**, 1944–1947 (2004). **This paper showed how addiction can be conceptualized as a disease of the habit valuation system, using a simple modification of the reinforcement-learning model.**
133. Paulus, M. P. Decision-making dysfunctions in psychiatry—altered homeostatic processing? *Science* **318**, 602–606 (2007).
134. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
135. Hazy, T. E., Frank, M. J. & O'Reilly, R. C. Towards an executive without a homunculus: computational models of the prefrontal cortex/basal ganglia system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 1601–1613 (2007).
136. Niv, Y., Joel, D. & Dayan, P. A normative perspective on motivation. *Trends Cogn. Sci.* **10**, 375–381 (2006).

Acknowledgements

Financial support from the National Science Foundation (SES-0134618, A.R.) and the Human Frontier Science Program (C.F.C.) is gratefully acknowledged. This work was also supported by a grant from the Gordon and Betty Moore Foundation to the Caltech Brain Imaging Center (A.R., C.F.C.). R.M. acknowledges support from the National Institute on Drug Abuse, the National Institute of Neurological Disorders and Stroke, The Kane Family Foundation, The Angel Williamson Imaging Center and The Dana Foundation.

FURTHER INFORMATION

Author's homepage: www.rnl.caltech.edu

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

Neuromarketing: the hope and hype of neuroimaging in business

Dan Ariely and Gregory S. Berns

Abstract | The application of neuroimaging methods to product marketing — neuromarketing — has recently gained considerable popularity. We propose that there are two main reasons for this trend. First, the possibility that neuroimaging will become cheaper and faster than other marketing methods; and second, the hope that neuroimaging will provide marketers with information that is not obtainable through conventional marketing methods. Although neuroimaging is unlikely to be cheaper than other tools in the near future, there is growing evidence that it may provide hidden information about the consumer experience. The most promising application of neuroimaging methods to marketing may come before a product is even released — when it is just an idea being developed.

Despite many common beliefs about the inherently evil nature of marketing, the main objective of marketing is to help match products with people. Marketing serves the dual goals of guiding the design and presentation of products such that they are more compatible with consumer preferences and facilitating the choice process for the consumer. Marketers achieve these goals by providing product designers with information about what consumers value and want before a product is created. After a product emerges on the marketplace, marketers attempt to maximize sales by guiding the menu of offerings, choices, pricing, advertising and promotions.

In their attempts to provide these types of inputs, marketers use a range of market research techniques, from focus groups and individual surveys to actual market tests — with many approaches in between (see *Supplementary information S1* (box)). In general, the simpler approaches (focus groups and surveys) are easy and cheap to implement but they provide data that can include biases, and are therefore seen as not very accurate^{1–4}. The approaches that are more complex and therefore harder to implement, such as market tests, provide more accurate data but incur a higher cost, and the product, production and distribution systems have to be in place for market tests to be conducted. There are some compromise approaches between these two extremes, which include simulated markets, conjoint analyses, markets for information and incentive-compatible pricing studies (see *Supplementary information S1* (box)).

As in all compromises, these approaches provide solutions with intermediate levels of cost, simplicity, realism and quality of data (TABLE 1).

The incorporation of neuroimaging into the decision-making sciences — for example, neuroeconomics — has spread to the realm of marketing. As a result, there are high hopes that neuroimaging technology could solve some of the problems that marketers face. A prominent hope is that neuroimaging will both streamline marketing processes and save money. Another hope is that neuroimaging will reveal information about consumer preferences that is unobtainable through conventional methods. Of course, with such high expectations, there is the accompanying hype. Several popular books and articles have been published that push a neuromarketing agenda, and there are now a handful of companies that market neuromarketing itself⁵. In this Perspective, we aim to distinguish the legitimate hopes from the marketing hype. As such, we hope that this article serves the dual purpose of recognizing the real potential of neuroimaging in business and providing a guide for potential buyers and sellers of such services.

Why use brain imaging for marketing?

Marketers are excited about brain imaging for two main reasons. First, marketers hope that neuroimaging will provide a more efficient trade-off between costs and benefits. This hope is based on the assumptions that people cannot fully articulate their preferences when asked to express them explicitly, and that consumers' brains contain hidden

information about their true preferences. Such hidden information could, in theory, be used to influence their buying behaviour, so that the cost of performing neuroimaging studies would be outweighed by the benefit of improved product design and increased sales. In theory, at least, brain imaging could illuminate not only what people like, but also what they will buy.

Thus far, this approach to neuromarketing has focused on this post-design application, in particular on measuring the effectiveness of advertising campaigns. The general approach has been to show participants a product advertisement, either in the form of a print advertisement or commercial, and measure the brain's response in the form of a blood oxygenation level-dependent (BOLD) measurement, which is taken as a proxy for neural activation.

The second reason why marketers are excited about brain imaging is that they hope it will provide an accurate marketing research method that can be implemented even before a product exists (FIG. 1). The assumption is that neuroimaging data would give a more accurate indication of the underlying preferences than data from standard market research studies and would remain insensitive to the types of biases that are often a hallmark of subjective approaches to valuations. If this is indeed the case, product concepts could be tested rapidly, and those that are not promising eliminated early in the process. This would allow more efficient allocation of resources to develop only promising products.

Thus, the issue of whether neuroimaging can play a useful part in any aspect of marketing depends on three fundamental questions, which we will address in this paper. First, can neuromarketing reveal hidden information that is not apparent in other approaches? Second, can neuromarketing provide a more efficient cost-benefit trade-off than other marketing research approaches? Third, can neuromarketing provide early information about product design?

Revealing hidden information

Brain activity and preference measurement.

Allowing for the assumption in neuromarketing that the brain contains hidden information about preferences, it is reasonable to set aside, for the moment, the issue of 'hidden' and ask what relationships are known to exist between brain activity and expressed (that is, not hidden) preference.

As it turns out, different methods of eliciting a person's preference often result in different estimations of that preference^{3,4,6,7}.

Table 1 | Comparison of selected marketing research approaches

	Focus groups	Preference questionnaires	Simulated choice methods	Market tests
What is measured	Open-ended answers, body language and behaviour; not suitable for statistical analysis	Importance weighting for various product attributes	Choices among products	Decision to buy and choice among products
Type of response process	Speculative, except when used to assess prototypes	The respondent must try to determine his decision weightings through introspection, then map those weightings into the response scale	A hypothetical choice, so the same process as the actual purchase — but without monetary consequences	An actual choice, with customers' own money, and therefore fully consequential
Typical use in new-product development processes	Early on to aid general product design; at user interface design for usability studies	Design phase, when determining customer trade-offs is important	Design phase, when determining customer trade-offs is important; may also be used as a forecasting tool	End of process, to forecast sales and measure the response to other elements of marketing, such as price
Cost and competitive risk	Low cost; risk comes only from misuse of data by the seller	Moderate cost and some risk of alerting competitors	Moderate cost (higher if using prototypes instead of descriptions) and some risk of alerting competitors	High cost and high risk of alerting competitors, plus the risk of the product being reverse engineered before launch
Technical skill required	Moderation skills for inside the group and ethnographic skills for observers and analysts	Questionnaire design and statistical analysis	Experiment design and statistical analysis (including choice modelling)	Running an instrumented market and forecasting (highly specialized)

This makes it difficult to know which method provides the truest measure of 'decision utility' (that is, the expected utility, which would ultimately drive choice in the marketplace). It is clear that market tests give the most accurate answer, but having to run a market test on every product would defeat the purpose of market research — namely, to provide early and cheap information. Similarly, we suspect (and economists are certain) that methods that are incentive compatible are better than methods that are not. Incentive-compatible elicitation methods are methods that encourage the participant to truthfully reveal what is being asked of him because to do so would maximize the participant's satisfaction (for example, he would earn the most money or receive the product he likes the best). In other words, it is in the participant's interest to answer product-related questions truthfully. However, using such methods is not always possible.

One important question for the potential of neuromarketing is whether the neural signal at the time of, or slightly before, the decision (assumed to be a measure of decision utility) can be a good predictor of the pleasure or reward at the time of consumption (the 'experienced utility')⁸. A second question is whether the link between these two signals holds even when the preference elicitation methods are not incentive compatible. If the answer

to both of these questions is positive, neuromarketing could become useful for measuring preferences.

Measurements such as willingness to pay (WTP) have only recently come under functional MRI (fMRI) examination. In one experiment, subjects bid on the right to eat snacks during the experiment. The amount they were willing to pay (a measure of decision utility) correlated with activity levels in the medial orbitofrontal cortex (OFC) and prefrontal cortex (PFC)^{9,10}. Interestingly, similar activation in the OFC has been observed when subjects anticipate a pleasant taste¹¹, look at pretty faces¹², hear pleasant music¹³, receive money^{14,15} and experience a social reward^{16,17}. Such generally close correspondence in regional brain activity between the anticipation of rewarding events, the consumption of enjoyable goods and the willingness to pay for them suggests that the representation of expected utility may rely, in part, on the systems that evaluate the quality of the consumption experience. The theme of common systems for expectation and experience also applies to things that are unpleasant or even painful (although this involves a different network including the insula)^{18–21}. Such similarities suggest that neuroimaging can become a useful tool in measuring preferences, particularly when incentive compatibility is important but there is no easy way to achieve it (for example, when the products have not been created).

However, such similarities do not necessarily mean that brain activation is the same across different elicitation methods, and there are differences between the neural activation representing decision utility and that representing experienced utility^{14,22,23}. This caveat aside, the generally close correspondence does suggest that neural activity might be used as a proxy for WTP in situations in which WTP cannot easily be determined — although this has yet to be demonstrated.

Reverse inference and reward. The practice of measuring an increase in BOLD activity in a region such as the ventral striatum or OFC and then concluding that a 'reward-related' process was active has become increasingly common. This form of deductive reasoning is known as 'reverse inference'^{24,25}. Given the readiness of many to interpret brain activation as evidence of a specific mental process, it is worth examining this type of inference. Using a Bayesian analysis, it is possible to estimate the specificity of activation in a particular region of the brain for a specific cognitive process. For example, Poldrack used the BrainMap database to analyse the frequency of activation of Broca's area in language studies²⁴. He found that activation of Broca's area implied a Bayes factor of 2.3 for language involvement, which means that taking brain activity into account can make a small but significant

improvement to one's prior estimate of whether a language process was involved.

Many studies have shown that striatal activity correlates with hedonic rating scales²⁶. Neuromarketers have been quick to invert this finding and use ventral striatal activity as an indication that an individual likes something; but what is the evidence for this? Using Poldrack's method to analyse the BrainMap database, we estimated the posterior probability for a reward process given the observation of nucleus accumbens (NAc) activation²⁷. The prior probability of engaging a reward-related process was assumed to be 0.5 (1:1 odds). According to this estimation, based on the number of fMRI papers reported in the BrainMap database with and without 'reward' and with and without NAc activation, NAc activation increases the probability of a reward-related process taking place to 0.90 (odds 9:1). This yields a Bayes factor of 9, which is considered moderate to strong evidence for a causal relationship (BOX 1). Although meaningful in a statistical sense, the assumptions behind such a calculation are rather liberal and may suffer from a publication bias for positive results as well as differing definitions of reward. In real-world settings, the ability to infer whether an individual likes something based on NAc activation alone may be substantially less.

In the context of a product likeability experiment, Knutson *et al.* found significant correlations between NAc activity and product preferences in college students²⁸. However, in logistic regression (R^2) calculations aimed at predicting consumer choice, self-reported preferences outperformed brain activation alone. Adding brain activation to a logistic model improved predictions, but only slightly (increasing R^2 from 0.528 to 0.533). Re-analysis with more sophisticated machine-learning algorithms further improved the predictive value of brain activation²⁹.

Although some have argued for the existence of a "buy button" in the brain⁵, current evidence suggests that the cognitive processes associated with purchase decisions are multifactorial and cannot be reduced to a single area of activation. Conversely, a given brain region may be involved in multiple cognitive processes. A recent review of value-based decision making divided the process of making a choice into five categories: representation of the decision; assignment of value to different actions; action selection; outcome evaluation; and learning³⁰. Even within this simplified framework, current data suggest that responses to marketing

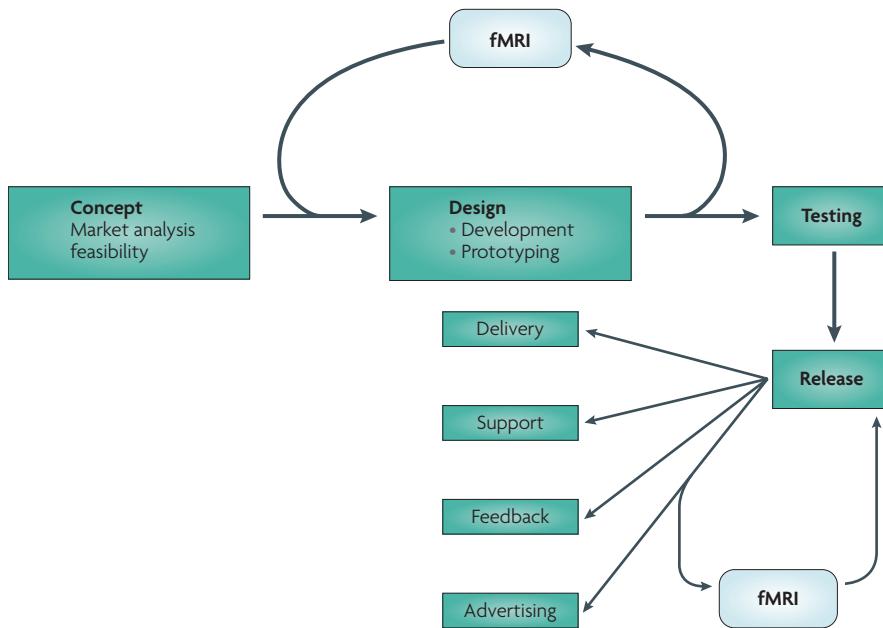


Figure 1 | Product development cycle. Neuromarketing applications of functional MRI (fMRI) can potentially enter into the product development cycle in two places. In the first, fMRI can be used as part of the design process itself. Here, neural responses could be used to refine the product before it is released. In the second, fMRI can be used after the product is fully designed, typically to measure neural responses as part of an advertising campaign to increase sales.

efforts and consumer choices depend on an array of neurobiological processes, and that no single brain region is responsible for a consumer choice. But is it possible that some brain regions are more involved than others? Because the field of neuroeconomics grew out of early brain-imaging studies of the neurobiology of reward^{31,32}, most of the neuroeconomic data are about valuation mechanisms and the associated responses of dopamine-rich brain regions. The OFC and striatum have been consistently implicated in goal-directed action^{9,22,33–35}. It is also generally accepted that the insula has a key role in physiological arousal, which is typically, although not exclusively, aversive in nature²¹. But because of the reverse inference problem, using striatal and OFC activity as a read-out of 'liking' and the insula as a 'disgust-meter' is probably too simplistic to be of use in a real-life setting. In the context of neuromarketing, the statistical power of these single-region correlations may be too low for the correlations to be of use as predictors of consumption unless, perhaps, the neuroimaging data is combined with other measures of preference.

fMRI as a brain decoder. Given the limited power of reverse inference from single-region brain activations, more data-driven methods for interpreting brain imaging

data have been at the forefront of analysis techniques. These techniques treat sites of brain activity agnostically — that is, without reference to prior hypotheses. The primary assumption is that, regardless of how an individual's brain represents information, it does so consistently. The representations may be spatially dispersed, and they may be distributed differently in different individuals, but they can still be reliably detected through multi-voxel pattern analysis (MVPA). Because MVPA methods are not reliant on the activation of a small subset of brain regions, they have substantially increased sensitivity to detect activation³⁶. A crucial advantage of MVPA techniques over approaches in which activation in a particular brain region of interest is measured is that MVPA has the statistical power to predict the individual choices of a subject. Because MVPA involves statistical associations of complex activation patterns that occur when an individual choice is being made, it does not depend on the vagaries of an experimenter interpreting the meaning of an activation map. Some of the most impressive demonstrations of MVPA have been in decoding visual responses to simple stimuli^{37–39} and subsequently, to watching films⁴⁰, the meanings of nouns⁴¹, event boundaries of written narratives⁴² and city navigation^{43,44}.

Box 1 | NAc activation in studies of tasks with and without reward

The BrainMap database was searched for functional MRI studies with and without a reward task and with and without nucleus accumbens (NAc) activation. The NAc was defined as a bilateral region of interest with vertices from MNI (Montreal Neurological Institute) coordinates ($-12, 0, -12$) to ($12, 12, 0$). The frequencies that were obtained are shown in the table below.

Assuming that the prior probability of engaging in a reward-related process is 0.5, calculations showed that NAc activation increases the probability of a reward-related process taking place to 0.90, yielding a Bayes factor of 9:

$$\text{Probability of NAc activation given a reward task} = 27/68 = 0.397$$

$$\text{Probability of NAc activation given no reward task} = 59/1283 = 0.046$$

Assuming the prior probability of reward = 0.5, then

$$\text{Probability of a reward task given NAc activation} = \frac{0.397}{(0.397 + 0.046)} = 0.90$$

	Reward task	No reward task
NAc activated	27	59
NAc not activated	41	1,224

It is possible, even likely, that such methods will soon be able to handle almost any circumstance that can be created in an MRI environment. With increasing stimulus complexity, simple interpretations of brain activation will become more difficult. However, for real-world marketing applications, it may be more important to predict future behaviour than to understand the 'why' of behaviour. Such a data-driven application of imaging (perhaps even lacking an underlying theory) is analogous to identifying a genetic polymorphism associated with a particular cancer without understanding what that gene does — which is likely to yield specific but not general insights.

Costs and benefits

As noted above, it is not yet clear whether neuroimaging provides better data than other marketing methods (TABLE 1), but through the use of MVPA methods it might be possible to reveal the 'holy grail' of hidden information. Assuming that this is the case, will using expensive neuroimaging ultimately be more efficient than using cheaper methods? Typical charges for scanning in a university research setting average about US\$500 per hour. In a commercial setting, these will be higher. However, actual scan charges account for a small portion of the total cost, with personnel and overhead expenses accounting for at least 75% of the costs of an imaging project. If neuromarketing is to compete with conventional marketing approaches on the basis of efficiency, then the costs of labour and overheads will have to be reduced.

One area in which the cost of neuro-imaging can be compared with conventional

marketing approaches is in the post-design phase, the goal of which is to increase sales of an existing product — for example, through advertisements and other types of framing effects. Early neuromarketing studies therefore used imaging approaches to evaluate consumer responses to advertisements. At this point, it is important to distinguish between neural responses to the consumption of a product (that is, experienced utility) and neural responses to representations of the product that may lead to future consumption. Only certain types of products can be consumed in an MRI scanner. Therefore, much of the post-design neuromarketing literature has focused on brain responses to visual representations of products, such as pictures^{28,45} or advertisements for the product^{46–48}; however, these advertisement studies, which used magnetoencephalography and electroencephalography (BOX 2), did not link imaging data to actual purchase decisions or other ratings, so it is not yet possible to determine the value of this approach.

The role of expectations. It has long been known that the manner in which choices are presented can have a dramatic effect on decisions⁴⁹. This is where advertisements and product placement come into play. To date, experiments have examined fairly simple choices and responses to things that can be presented in an MRI scanner. Before neuroimaging can be used to predict consumer choice, a greater understanding of the interplay between the decision maker, the elicitation method and the decision context is needed.

BOLD responses are influenced by so-called 'expectation' effects, which include pricing effects, biases in the way the choice is presented⁵⁰ and placebo responses. This suggests that neuromarketing could be helpful in identifying individual differences in consumer reactions to different types of inputs. In a study of neural responses to sips of wine, medial OFC responses were higher when subjects were told that the wine was expensive (\$90 per bottle) versus inexpensive (\$5 per bottle)²³. Activity in this region also correlated with self-report ratings of how much participants liked the wine, even though all wines were actually the same. These results suggest that the instantaneous experience of pleasure from a product — that is, experienced utility — is influenced by pricing, and that this effect may be mediated by the medial OFC⁹. This result parallels a similar, behavioural finding that the strength of the placebo effect for analgesia is greater for more expensive 'medications'⁵¹. Subjects' expectations also play an important part in how the experimenter should interpret striatal responses. Many studies have shown that the reward-related signals in the ventral striatum and NAc can be more accurately linked to prediction errors for reward than to reward itself^{22,52,53}.

Placebo responses are an interesting aspect of neuromarketing. The mechanism of the placebo response has been debated for decades⁵⁴, but ultimately it can be considered an effect of marketing (that is, the actions of a doctor, pharmaceutical company or experimenter). The neural correlates of the analgesic placebo effect are widespread but generally point to a modulation of the cortical pain matrix in the brain^{55,56}. Because consumers cannot consciously report placebo effects, the demonstration of neural correlates of these effects suggests that having access to hidden brain information could enable a marketer to measure the effectiveness of a placebo marketing strategy in a particular individual. How well this type of information generalizes to a larger population will determine the cost–benefit ratio of doing neuroimaging.

The aforementioned manipulations of expectations are simple and direct. For example, the experimenter can manipulate a single dimension of expectation, such as price or descriptive words (for example, "ultra" and "new and improved"), and measure the effect on the consumer behaviourally and neurally. More cognitively complex forms of expectations can be created through advertisements and commercials.

Box 2 | Neuromarketing technologies

Functional MRI (fMRI)

The technique uses an MRI scanner to measure the blood oxygenation level-dependent (BOLD) signal. The BOLD changes are generally correlated with the underlying synaptic activity. Spatial resolution is 1–10 mm, and temporal resolution is 1–10 s. In general, the higher the spatial resolution, the lower the temporal resolution. Of the three imaging technologies described in this Box, fMRI has a substantial advantage in resolving small structures and those that are deep in the brain. However, some important brain regions, especially the orbitofrontal cortex, are affected by signal artefacts that may reduce the ability to obtain useful information. State-of-the-art MRI scanners cost approximately US\$1 million per Tesla and have annual operating costs of \$100,000–\$300,000.

Electroencephalography (EEG)

EEG uses electrodes applied to the scalp and measures changes in the electrical field in the brain region underneath. EEG has very high temporal resolution (milliseconds) and can therefore detect brief neuronal events. Because the skull disperses the electrical field, EEG has low spatial resolution (~1 cm) that depends on how many electrodes are used. The number of electrodes can be as few as two or range up to hundreds in high-density arrays. The greater the number of electrodes, the better the spatial resolution. Apart from the low spatial resolution, EEG has poor sensitivity for deep brain structures. Equipment costs can be low (<\$10,000) but increase with high-density arrays and the concomitant resources needed to process the data. A common technique is to measure the left-right asymmetry of the frontal EEG⁷⁸. This is typically measured by the power in the alpha band (8–13 Hz). This research has suggested that relatively greater activity in the left frontal region is associated with either positive emotional states or the motivational drive to approach an object⁷⁹. Although there are strong correlations between frontal EEG asymmetry and personality traits, the degree to which the asymmetry changes from moment to moment is still debated. Some have suggested a minimum of 60 s to reliably estimate power asymmetry⁸⁰, in which case the temporal advantage of EEG over fMRI is lost. Although some have used this approach to measure momentary fluctuations in emotion in response to advertisements⁸¹, without accounting for autocorrelations in time or multiple statistical comparisons, the validity of such approaches is dubious.

Magnetoencephalography (MEG)

An expensive cousin of EEG, MEG measures changes in the magnetic fields induced by neuronal activity. Thus, MEG has the same advantage of high temporal resolution and, because the magnetic field is less distorted by the skull than is the electrical field, it has better spatial resolution than EEG. Like EEG, MEG is most sensitive to superficial cortical signals (primarily in the sulci). MEG requires a magnetically shielded room and superconducting quantum interference detectors to measure the weak magnetic signals in the brain. An MEG set-up costs approximately \$2 million.

Transcranial magnetic stimulation (TMS)

TMS uses an iron core, often in the shape of a toroid wrapped in electrical wire, to create a magnetic field strong enough to induce electrical currents in underlying neurons when placed on the head⁸². TMS can be used as a single pulse, paired pulse or repetitive stimulation, and the neuronal effects range from facilitation to inhibition of synaptic transmission. As a research tool, TMS has been used to study the causal role of specific brain regions in particular tasks by temporarily taking them 'offline'.

Post-design applications of neuroimaging have, for the most part, confirmed what was known about the behavioural effects of product placement, which bypass the counter-arguments in which people naturally engage when facing advertisements. The imaging studies confirm that there are neural correlates of exposure to advertisements but do not directly suggest that maximizing activity in a particular brain region results in more sales.

Culture and advertising. Neuroimaging is often hyped as an exciting new tool for advertisers. Despite its enormous cost, advertising effectiveness is a poorly understood area of marketing. Although advertising has been investigated in a few

neuroimaging studies^{57,58}, it is still unknown whether neuroimaging can prospectively reveal whether an advertisement will be effective. In a famous Coke–Pepsi study, participants who described themselves as Coke drinkers showed significant activation in the hippocampus and right dorsolateral PFC when they were cued about the upcoming drink of Coke⁴⁵. Self-described Pepsi drinkers did not have this response. In the absence of brand information, there was no significant difference in preference during a taste test. The study suggested that any differences in the response (behavioural and neural) to the two brands must be culturally derived. One possibility is that brands achieve a life of their own by becoming animate objects, sometimes with human attributes, in the

minds of consumers. However, one fMRI study that compared brain responses to persons and brands found that activation patterns for brands differed from those for people — even for brands with which subjects are identified — suggesting that brands are not perceived in the same way as people⁵⁹. Another possibility is that specific emotions can be elicited in response to advertisements, although whether neuroimaging will help to reveal these emotions may ultimately be limited by reverse inference constraints, especially if tied to specific regions.

The issue of how culturally derived identities become embedded in the brain is of great interest, not only from a marketing perspective. Although neoclassical economic theory describes a framework in which individuals assess costs and benefits during their decision-making processes, it is clear that people base many decisions on sociocultural rules and identities. Some are in a commercial context (for example, "I am a PC" or "I am a Mac") but many are not (for example, "I am a Democrat" or "I am a Republican"). These issues extend beyond the mundane questions of advertisement effectiveness and raise the more profound question of how the marketing of ideas affects decision making. But whether neuroimaging provides an efficient tool to answer this question has yet to be shown.

Early product design

As the ability of neuroimaging to predict or influence post-design purchase decisions seems to be limited (see above), neuroimaging may be better suited to gauging responses before products are marketed. The primary reason is that neuroimaging may yield insights into the product experience itself.

Food products. Various food products and beverages have been administered in the MRI scanner, from simple sugar solutions to chocolate, wine, sports drinks and colas. Beverages are particularly easy to administer, with the usual route through a computer-controlled pump attached to a tube that delivers controlled amounts of fluid into the participant's mouth. The perception of flavour is a multisensory integration process and thus provides several opportunities for neuroimaging to disentangle a complex perception that subjects might not be able to articulate; taste, odour, texture, appearance and even sound all contribute to the gustatory experience. These different dimensions have been mapped onto

Box 3 | The ethics of neuromarketing

The introduction of neuroimaging into an environment in which the ultimate goal is to sell more product to the consumer may raise ethical issues.

- **Businesses will be able to read the minds of consumers.** This concern is about the privacy of thoughts. Can neuroimaging be used to gauge a person's preferences outside of the specific task being performed? Possibly. This concern may be mitigated through transparency of purpose: subjects must know what kind of endeavour they are helping, and their data should only be used for that purpose.
- **Private versus public information about preferences.** Individuals need to be able to exercise control over what they choose to reveal about their personal preferences. A privacy breach occurs if neuroimaging reveals a private preference that is outside the scope of the neuromarketer's research question.
- **Information will be used to discriminate against individuals or exploit particular neurological traits found in a subgroup of individuals.** Neuroimaging data could potentially target marketing to specific people or groups. Many people would find this tactic repugnant because it exploits a biological 'weakness' that only exists in some people. Similarly, this information could be used to time pricing moves to capitalize on individual weaknesses that are known to coincide with particular biological states (for example, raising beverage prices when someone is known to be thirsty).
- **Central versus peripheral routes of influence.** A central route aims to influence consumers' preferences about the functional aspects of the product (for example, fewer calories in a beer). A peripheral route attempts to manipulate preferences through things that are peripherally related to the product (for example, sex appeal of people in advertisements). Neuroimaging could potentially be used to enhance both types of influence, but some consider the attempts to optimize the peripheral route more ethically dubious.
- **Brain responses obtained from a small group of subjects will be used to generalize to a large population.** Of course, this is done all the time in the scientific literature. If neuromarketing data are used in product design and the product injures someone, neuroimaging will be partly to blame.
- **Abnormal findings.** Approximately 1% of the population will have an abnormality on their MRI⁶³. In a population without clinical symptoms, the clinical significance of an MRI abnormality is unknown. Many will be false positives; others will be real and require referral. Currently, there is no standard for how to handle these situations. However, it is standard practice to have a written policy in place for abnormal findings. Failure to do so opens both the neuromarketing firm and their clients to medical liability.
- **A lack of regulation.** Traditional marketing methods, because they are not typically viewed as experimentation, have not been subject to institutional review board (IRB) oversight. MRI scans are approved by the US Food and Drug Administration (FDA) for clinical use but, because no diagnosis is being made in a marketing setting, there is the potential to circumvent both FDA and IRB requirements. The burgeoning neuromarketing industry would be well advised to adopt an industry standard of independent review. Clients should demand it.
- **Management of perceptions.** How will the public react when they discover that neuroimaging has been used to design or market a product? The public's response to genetically modified food could provide an indication.
- **Companies might not be primarily concerned with the best interests of the consumer.** Companies and consumers maintain complex relationships in which some of their goals are compatible while others are in conflict. On the one hand, companies seek to design, manufacture and sell products that consumers seek to buy, resulting in compatible goals that benefit both parties. On the other hand, companies also aim to maximize their short- or long-term profits, sometimes to the detriment of their consumers. Much like marketing itself, understanding consumer preferences can be used for goals that are in the best interests of both the company and their consumers or for objectives that are in the interests of the company and to the detriment of their consumers. Which approaches neuromarketers choose is an open question.

distinct brain regions but with substantial overlap^{60,61}. The OFC is consistently linked to perceived pleasantness, whereas viscosity and fat content seem to be represented in the insula⁶². The use of neuroimaging by commercial manufacturers to design a more appealing food product is both feasible and likely. For this to work, however, one would need to identify which

dimension of gustation is to be studied (for example, taste, odour or texture) and maximize a brain response to variations in that dimension.

The drawback to such an approach is the possibility of creating food products that are so highly tuned to neural responses that individuals may over-eat and become obese (see BOX 3 for a discussion of some ethical

issues related to neuromarketing). Is it possible that such a neuroimaging approach could create a 'super-heroin of food' — a product so delicious that all but the most ascetic individuals would find it irresistible? It is an extreme but real possibility. However, that does not mean that neuroimaging is necessarily problematic for food product development. Indeed, the same techniques could be applied to making nutritious foods more appealing.

Entertainment. As a typical big-budget Hollywood film costs over \$100 million, with almost as much spent on marketing, it would be surprising if film producers were not interested in using neuroimaging to improve their product. After static images, films are probably the easiest product to present in the scanner. Moreover, an fMRI measurement is time locked to the film timeline. A film presents the same basic visual and auditory stimuli to everyone viewing it and thus should serve as a cognitive synchronizer. Indeed, an fMRI study of subjects viewing a segment of the classic Western *The Good, the Bad and the Ugly*⁴⁰ showed that large extents of the cortex responded similarly in time across subjects, suggesting that much of the cortical response is essentially stereotypical. In another study, the ability to recall narrative content of the TV sitcom *Curb Your Enthusiasm* three weeks later was correlated with the strength of hippocampal and temporal lobe responses during viewing⁶³.

Such stereotypical responses suggest that fMRI could be used during the editing process. For example, different cuts of a movie could be measured against these cortical responses, which could then be used to select the final cut for release. Although it seems hopelessly complex to interpret such brain responses, it may not be necessary if the only goal is to release the most profitable movie. Provided there were a metric of quality (for example, box office returns or test audience reports), brain activation patterns could be chosen to optimize outcomes without any knowledge of what the patterns meant. Several neuromarketing companies have targeted their efforts towards the entertainment industry but, as most of this work is unpublished, it is difficult to evaluate the quality of the product. However, guidelines for general quality of scientific work can be formulated based on two decades of neuroscience research. Thus, without passing judgment on whether neuromarketing works, we can at least identify the items to look for in a quality operation (BOX 4).

Box 4 | What to look for when hiring a neuromarketing firm

We provide a list, which is by no means exhaustive, of what could be considered standard practice in the application of neuroimaging methods in cognitive neuroscience and related fields. It is based on standard criteria for reviewing research proposals and adapted to a business setting.

- What is to be gained from neuroimaging? Good neuromarketers will begin by discussing the pros and cons of the proposal in detail. For example: what will neuroimaging yield over traditional methods? Ask for data about the predictive value of neuroimaging findings in a real-world setting.
- What are the dependent and independent measurements? Assessing brain activation is not generally useful without correlating it with some other measurement. It is necessary to have another behavioural measurement to anchor the interpretation of the brain activation. Be wary if someone claims to know what a person thinks based solely on brain activation.
- How many subjects are needed? Apart from the simplest of tasks, any task invoking a response that is expected to vary across individuals demands a sample size of at least 30 (REF. 84). If groups of individuals are being compared under different treatments or conditions, the sample size will need to be much greater to detect differences between groups and between different treatments.
- What is the nature of the stimuli? Simple stimuli are the easiest to analyse. Real-world images, as might appear in an advertisement, become difficult to characterize unless one element at a time is varied. For statistical power, a minimum of 10 repetitions within a stimulus category are required, although 20–30 would be more likely to achieve meaningful results.
- What type of software will be used to analyse the neuroimaging data? Several software packages exist, and although these programmes make neuroimaging seem simple, it takes a minimum of 1 year of training to be able to use them and 3 years to become fully competent.
- How will motion correction be performed?
- Are conditions balanced in time? If not, how will subjects' drifting attention be compensated for?
- Is this a whole-brain analysis or is a specific part of the brain being examined? These necessitate different thresholds of identifying activation. The chance of an activation appearing somewhere in the brain is high due to random noise.
- Will regions of interest be defined *a priori*? If so, what is the justification for this? Conclusions based on activation of a single region will have relatively little predictive power over conventional behavioural methods.
- If multi-voxel pattern analysis (MVPA) methods will be used, will they be completely data-driven (principal component analysis or independent component analyses) or will they be based on classifier training of subject responses (support vector machine, relevance vector regression or Gaussian process regression)? How will the resulting activity maps be interpreted?
- How robust are the results? Ask for a 'bootstrap' — for example, testing on a 'fresh' subsample of data.
- What type of scanner will be used? Either 1.5 or 3 Tesla scanners can yield images of acceptable quality. Open MRIs do not have the field homogeneity or the gradient technology necessary for fMRI. What quality control checks are performed to make sure the scanner is operating optimally and consistently from day to day? What steps will be taken to minimize signal artefacts in areas with poor signal?

Architecture. A growing number of neuroscientists and architects have begun to consider the relationships of the brain to the architectural experience⁶⁴. The neuroscience of architecture could be considered from two perspectives: first, the neural activity associated with seeing specific aspects of a building; and second, the use of neural responses to guide the architectural design process. Clearly, one would need to identify these neural responses before attempting to use them in architectural design, but it is precisely the application in design that places neuroimaging within the neuromarketing framework.

Virtual reality can provide a surprisingly accurate simulation of an architectural experience and can be used in an MRI scanner. It

has already been used to understand neural activation during automobile driving^{65,66}. In spatial navigation tasks such as driving, and presumably navigating a building, the hippocampus has a key role. These early virtual reality experiments suggested that the hippocampus is active when the subject makes navigation decisions but not when they are externally cued⁶⁵. Perhaps taking into account 'hippocampal load' may be a useful tool in architectural design — for example, to make buildings easier to navigate. Extending this idea by considering the neurobiological changes associated with ageing, it might be possible to design buildings and retirement communities that mitigate the memory loss associated with Alzheimer's disease.

Political candidates. Finally, neuromarketing might be applied to perhaps the greatest marketing campaign of all: politics. According to the [Federal Election Commission](#) (see Further information), the cost of the 2008 US Presidential race was approximately \$1.6 billion. It was also around that time that neuroimaging made its way into politics, perhaps most prominently in the form of a *New York Times* op-ed piece⁶⁷. Peer-reviewed studies have shown a complex pattern of activation in response to statements about candidates; these patterns have been interpreted as evidence that motivated reasoning involves activation in the ventromedial PFC, the anterior cingulate cortex, the posterior cingulate cortex and the insula⁶⁸. Subsequent studies have suggested that activation of the medial PFC might be associated with maintaining a subject's preference for a candidate in response to advertisements, whereas activity in the lateral PFC might be associated with changing candidates⁶⁹.

In marketing terms, the political candidates are the products that must be sold to the electorate. Therefore, like other products, candidates and their campaigns have pre- and post-design phases. Political marketing is aimed at selling an existing candidate but, with more foresight, can also be used to 'design' a better candidate. The aforementioned neuroimaging studies have focused on the post-design responses to advertisements for political candidates^{68,69}.

Could neuroimaging also be used to design a candidate? Although potential nominees already go through a 'grooming' process, it is worth examining this prospect. A candidate's appearance, trustworthiness and message content might determine a voter's decision. Considerable neuroimaging work has been done on the perception of human faces⁷⁰ and features such as facial symmetry, skin colour and attractiveness. Key brain structures in visual processing include the fusiform face area for basic face processing⁷¹, the superior temporal sulcus for gaze direction and intention and the NAc for attractiveness¹². A recent study on the effect of political candidates' appearance found that insula activation in response to seeing a picture of a candidate was associated with a greater likelihood of that candidate losing the election⁷². In addition, dorsolateral PFC and anterior cingulate cortex activation occurred when subjects viewed images of a candidate of a political party different from their own⁷³. The neurobiology of trust has also become quite popular to study with both fMRI and, more recently, pharmacological

manipulations^{74–76}. These studies have found that different dimensions of trust, such as reputation, fairness and uncertainty, correlate with activity in different brain regions. Moreover, the hormone oxytocin affects human behaviour in various economic exchanges that depend on social interactions⁷⁷. Finally, a candidate's message content could be viewed as an experiential product. One could theoretically attempt to maximize striatal and OFC responses to platform statements although, for the reasons stated above, this is not necessarily predictive of success.

Conclusions and future directions

Neuromarketing has received considerable attention in both the scientific community and the media. Although few scientific neuromarketing studies have been conducted, the existing evidence suggests that neuroimaging could be used advantageously in several domains of marketing. For a marketer, neuroimaging could be attractive because it might be cheaper and faster than current marketing tools, and because it could provide hidden information about products that would otherwise be unobtainable. We think it unlikely that neuroimaging will be more cost-effective than traditional marketing tools, and so the first point is mostly hype. However, continuing developments in analytical tools for neuroimaging data — for example, MVPA — suggest that neuroimaging will soon be able to reveal hidden information about consumer preferences. Although this information could boost post-design sales efforts, we think that the real pay-off will come during the design process. Using fMRI data during design could affect a wide range of products, including food, entertainment, buildings and political candidates.

There are two sides to the use of such information. Product manufacturers could use neural information to coerce the public into consuming products that they neither need nor want. However, we hope that future uses of neuromarketing will help companies to identify new and exciting products that people want and find useful. One example is a new trend in 'user design' in which companies allow consumers to participate, through the internet, in the design of new products and by doing so create products that are more useful for the companies and for their customers. Perhaps a next phase in user design is one that incorporates not only what consumers express, but also what they think.

Finally, we return to the opening question: hope or hype? It is too early to tell but, optimists as we are, we think that there is much that neuromarketing can contribute to the interface between people and businesses and in doing so foster a more human-compatible design of the products around us. At the same time, neuromarketing as an enterprise runs the risk of quickly becoming yesterday's fad. Seasoned marketers still remember the hype around subliminal advertising, which quickly faded and died despite the research interest that surrounded it (and research on subliminal priming remains a large part of academic research in social psychology). How can we make sure that neuromarketing will not suffer a similar fate? For one, the academic community should take this topic seriously and not leave it to the neuromarketers and the op-ed page of the *New York Times*. We should also ask deeper questions on how marketing works — and not simply examine whether type X of advertising works better or worse than type Y. If we take neuromarketing as the examination of the neural activities that underlie the daily activities related to people, products and marketing, this could become a useful and interesting path for academic research and at the same time provide useful inputs to marketers.

Dan Ariely is at the Fuqua School of Business, Center for Cognitive Neuroscience, Department of Economics, and the Department of Psychiatry and Behavioural Sciences, Duke University, Durham, North Carolina 2770, USA.

Gregory S. Berns is at the Department of Psychiatry and Behavioural Sciences, Economics Department, Center for Neuropolicy, Emory University, Atlanta, Georgia 30322, USA.

*Correspondence to G.S.B.
e-mail: gberns@emory.edu*

doi: 10.1038/nrn2795

Published online 3 March 2010

- Beckwith, N. E. & Lehmann, D. R. The importance of halo effects in multi-attribute attitude models. *J. Mark. Res.* **12**, 265–275 (1975).
- Day, G. S. The threats to marketing research. *J. Mark. Res.* **12**, 462–467 (1975).
- Griffin, A. & Hauser, J. R. The voice of the customer. *Mark. Sci.* **12**, 1–27 (1993).
- Green, P. E. & Srinivasan, V. Conjoint analysis in marketing: new developments with implications for research and practice. *J. Mark.* **54**, 3–19 (1990).
- Lindstrom, M. *Buyology. Truth and Lies About Why We Buy* (Doubleday, New York, 2008).
- Hauser, J. R. & Shugan, S. M. Intensity measures of consumer preference. *Oper. Res.* **28**, 278–320 (1980).
- Buchanan, B. & Henderson, P. W. Assessing the bias of preference, detection, and identification measures of discrimination ability in product design. *Mark. Sci.* **11**, 64–75 (1992).
- Kahneman, D., Wakker, P. P. & Sarin, R. Back to Bentham? Explorations of experienced utility. *Q. J. Econ.* **112**, 375–405 (1997).
- Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* **27**, 9984–9988 (2007).
- Hare, T. A., Camerer, C. F. & Rangel, A. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science* **324**, 646–648 (2009).
- O'Doherty, J. P., Deichmann, R., Critchley, H. D. & Dolan, R. J. Neural responses during anticipation of a primary taste reward. *Neuron* **33**, 815–826 (2002).
- Aharon, I. *et al.* Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* **32**, 537–551 (2001).
- Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory-motor interactions in music perception and production. *Nature Rev. Neurosci.* **8**, 547–558 (2007).
- Knutson, B., Adams, C. M., Fong, G. W. & Hommer, D. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* **21**, RC159 (2001).
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J. & Andrews, C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neurosci.* **4**, 95–102 (2001).
- Izuma, K., Saito, D. N. & Sadato, N. Processing of social and monetary rewards in the human striatum. *Neuron* **58**, 284–294 (2008).
- Rilling, J. K. *et al.* A neural basis for social cooperation. *Neuron* **35**, 1–20 (2002).
- Ploghaus, A., Becerra, L., Borras, C. & Borsook, D. Neural circuitry underlying pain modulation: expectation, hypnosis, placebo. *Trends Cogn. Sci.* **7**, 197–200 (2003).
- Ploghaus, A. *et al.* Dissociating pain from its anticipation in the human brain. *Science* **284**, 1979–1981 (1999).
- Koyama, T., McHaffie, J. G., Laurienti, P. J. & Coghill, R. C. The subjective experience of pain: where expectations become reality. *Proc. Natl Acad. Sci. USA* **102**, 12950–12955 (2005).
- Craig, A. D. How do you feel? Interception: the sense of the physiological condition of the body. *Nature Rev. Neurosci.* **3**, 655–666 (2002).
- Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W. & Rangel, A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* **28**, 5623–5630 (2008).
- Plassmann, H., O'Doherty, J., Shiv, B. & Rangel, A. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl Acad. Sci. USA* **105**, 1050–1054 (2008).
- Poldrack, R. A. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* **10**, 59–63 (2006).
- Poldrack, R. A. The role of fMRI in cognitive neuroscience: where do we stand? *Curr. Opin. Neurobiol.* **18**, 223–227 (2008).
- Delgado, M. R. Reward-related responses in the human striatum. *Ann. NY Acad. Sci.* **1104**, 70–88 (2007).
- Fox, P. T. & Lancaster, J. L. Mapping context and content: the BrainMap model. *Nature Rev. Neurosci.* **3**, 319–321 (2002).
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D. & Loewenstein, G. Neural predictors of purchases. *Neuron* **53**, 147–156 (2007).
- Grosenick, L., Greer, S. & Knutson, B. Interpretable classifiers for fMRI improve prediction of purchases. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**, 539–548 (2008).
- Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nature Rev. Neurosci.* **9**, 545–556 (2008).
- Camerer, C., Loewenstein, G. & Prelec, D. Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* **43**, 9–64 (2005).
- Glimcher, P. W. Decisions, decisions, decisions: choosing a biological science of choice. *Neuron* **36**, 323–332 (2002).
- Yin, H. H., Ostlund, S. B. & Balleine, B. W. Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* **28**, 1437–1448 (2008).
- Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
- Schoenbaum, G. & Roesch, M. Orbitofrontal cortex, associative learning, and expectancies. *Neuron* **47**, 633–636 (2005).

36. Norman, K. A., Polyn, S. M., Detre, G. J. & Haxby, J. V. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* **10**, 424–430 (2006).
37. Haynes, J. D. & Rees, G. Decoding mental states from activity in humans. *Nature Rev. Neurosci.* **7**, 523–534 (2006).
38. Kamitani, Y. & Tong, F. Decoding the visual and subjective contents of the human brain. *Nature Neurosci.* **8**, 679–685 (2005).
39. Kay, K. N., Naselaris, T., Prenger, R. J. & Gallant, J. L. Identifying natural images from human brain activity. *Nature* **452**, 352–356 (2008).
40. Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
41. Mitchell, T. M. *et al.* Predicting human brain activity associated with the meanings of nouns. *Science* **320**, 1191–1195 (2008).
42. Speer, N. K., Zacks, J. M. & Reynolds, J. R. Human brain activity time-locked to narrative even boundaries. *Psychol. Sci.* **18**, 449–455 (2007).
43. Spiers, H. J. & Maguire, E. A. Spontaneous mentalizing during an interactive real world task: an fMRI study. *Neuropsychologia* **44**, 1674–1682 (2006).
44. Spiers, H. J. & Maguire, E. A. Decoding human brain activity during real-world experiences. *Trends Cogn. Sci.* **11**, 356–365 (2007).
45. McClure, S. M. *et al.* Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* **44**, 379–387 (2004).
46. Ambler, T., Ioannides, A. & Rose, S. Brands on the brain: neuro-images of advertising. *Bus. Strategy Rev.* **11**, 17–30 (2000).
47. Rossiter, J. R., Silberstein, R. B., Harris, P. G. & Nield, G. Brain-imaging detection of visual scene encoding in long-term memory for TV commercials. *J. Advert. Res.* **41**, 13–22 (2001).
48. Astolfi, L. *et al.* Neural basis for brain responses to TV commercials: a high-resolution EEG study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**, 522–531 (2008).
49. Tversky, A. & Kahneman, D. The framing of decisions and the psychology of choice. *Science* **211**, 453–458 (1981).
50. De Martino, B., Kumaran, D., Seymour, B. & Dolan, R. J. Frames, biases, and rational decision-making in the human brain. *Science* **313**, 684–687 (2006).
51. Waber, R. L., Shiv, B., Carmon, Z. & Ariely, D. Commercial features of placebo and therapeutic efficacy. *JAMA* **299**, 1016–1017 (2008).
52. Montague, P. R. & Berns, G. S. Neural economics and the biological substrates of valuation. *Neuron* **36**, 265–284 (2002).
53. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
54. Colloca, L. & Benedetti, F. Placebos and painkillers: is mind as real as matter? *Nature Rev. Neurosci.* **6**, 545–552 (2005).
55. Wager, T. D. *et al.* Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* **303**, 1162–1167 (2004).
56. Benedetti, F., Mayberg, H. S., Wager, T. D., Stohler, C. S. & Zubieta, J. K. Neurobiological mechanisms of the placebo effect. *J. Neurosci.* **25**, 10390–10402 (2005).
57. Kenning, P. H. & Plassmann, H. How neuroscience can inform consumer research. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**, 532–538 (2008).
58. Lee, N., Broderick, A. J. & Chamberlain, L. What is neuromarketing? A discussion and agenda for future research. *Int. J. Psychophysiol.* **63**, 199–204 (2007).
59. Yoon, C., Gutchess, A. H., Feinberg, F. & Polk, T. A. A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *J. Consum. Res.* **33**, 31–40 (2006).
60. Small, D. M. & Prescott, J. Odor/taste integration and the perception of flavor. *Exp. Brain Res.* **166**, 345–357 (2005).
61. Rolls, E. T. Brain mechanisms underlying flavour and appetite. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **361**, 1123–1136 (2006).
62. De Araujo, I. E. T. & Rolls, E. T. The representation in the human brain of food texture and oral fat. *J. Neurosci.* **24**, 3086–3093 (2004).
63. Hasson, U., Furman, O., Clark, D., Dudai, Y. & Davachi, L. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* **57**, 452–462 (2008).
64. Eberhard, J. P. Applying neuroscience to architecture. *Neuron* **62**, 753–756 (2009).
65. Spiers, H. J. & Maguire, E. A. Neural substrates of driving behaviour. *NeuroImage* **36**, 245–255 (2007).
66. Calhoun, V. D. *et al.* Different activation dynamics in multiple neural systems during simulated driving. *Hum. Brain Map.* **16**, 158–167 (2002).
67. Freedman, J. This is your brain on politics. *New York Times* (18 Jan 2005).
68. Westen, D., Blagov, P. S., Harenski, K., Kilts, C. & Hamann, S. Neural bases of motivated reasoning: an fMRI study of emotional constraints on partisan political judgment in the 2004 US Presidential election. *J. Cogn. Neurosci.* **18**, 1947–1958 (2006).
69. Kato, J. *et al.* Neural correlates of attitude change following positive and negative advertisements. *Front. Behav. Neurosci.* **3**, 6 (2009).
70. Tsao, D. Y. & Livingstone, M. S. Mechanisms of face perception. *Ann. Rev. Neurosci.* **31**, 411–437 (2008).
71. Kanwisher, N. & Yovel, G. The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Transact. Roy. Soc. B Biol. Sci.* **361**, 2109–2128 (2006).
72. Spezio, M. L. *et al.* A neural basis for the effect of candidate appearance on election outcomes. *Soc. Cogn. Affect. Neurosci.* **3**, 344–352 (2008).
73. Kaplan, J. T., Freedman, J. & Iacoboni, M. Us versus them: political attitudes and party affiliation influence neural response to faces of presidential candidates. *Neuropsychologia* **45**, 55–64 (2007).
74. Fehr, E. & Camerer, C. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* **11**, 419–427 (2007).
75. Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U. & Fehr, E. Oxytocin increases trust in humans. *Nature* **435**, 673–676 (2005).
76. King-Casas, B. *et al.* Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).
77. Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U. & Fehr, E. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* **58**, 639–650 (2008).
78. Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A. & Friesen, W. V. Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology I. *J. Pers. Soc. Psychol.* **58**, 330–341 (1990).
79. Harmon-Jones, E. Clarifying the emotive functions of asymmetric frontal cortical activity. *Psychophysiology* **40**, 838–848 (2003).
80. Huster, R. J., Stevens, S., Gerlach, A. L. & Rist, F. A spectroanalytic approach to emotional responses evoked through picture presentation. *Int. J. Psychophysiol.* **72**, 212–216 (2008).
81. Ohme, R., Reykowska, D., Wiener, D. & Choromanska, A. Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures. *J. Neurosci. Psychol. Econ.* **2**, 21–31 (2009).
82. Kobayashi, M. & Pascual-Leone, A. Transcranial magnetic stimulation in neurology. *Lancet Neurol.* **2**, 145–156 (2003).
83. Illes, J. *et al.* Incidental findings in brain imaging research. *Science* **311**, 783–784 (2006).
84. Thirion, B. *et al.* Analysis of a large fMRI cohort: statistical and methodological issues for group analyses. *NeuroImage* **35**, 105–120 (2007).

Acknowledgements

This work was supported by grants to G.S.B. from the National Institute on Drug Abuse (R01DA016434 and R01DA025045), the Office of Naval Research and Air Force Office of Scientific Research, and the National Science Foundation (BCS0827313).

Competing interests statement

The authors declare competing financial interests: see web version for details.

FURTHER INFORMATION

Dan Ariely's homepage:

<http://www.predictablyirrational.com/>

Gregory S. Berns's homepage:

<http://www.neuropolicyemory.edu/overview.html>

Federal Election Commission: <http://www.fec.gov/>

SUPPLEMENTARY INFORMATION

See online article: S1 (box)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF

Decision neuroscience and consumer decision making

**Carolyn Yoon · Richard Gonzalez ·
Antoine Bechara · Gregory S. Berns ·
Alain A. Dagher · Laurette Dubé · Scott A. Huettel ·
Joseph W. Kable · Israel Liberzon ·
Hilke Plassmann · Ale Smidts · Charles Spence**

Published online: 26 May 2012
© Springer Science+Business Media, LLC 2012

Abstract This article proposes that neuroscience can shape future theory and models in consumer decision making and suggests ways that neuroscience methods can be used in decision-making research. The article argues that neuroscience facilitates better theory development and empirical testing by considering the physiological context and the role of constructs such as hunger, stress, and social influence on consumer choice and preferences. Neuroscience can also provide new explanations

C. Yoon (✉)

Stephen M. Ross School of Business, University of Michigan, Ann Arbor, MI, USA
e-mail: yoenc@umich.edu

R. Gonzalez · I. Liberzon

University of Michigan, Ann Arbor, MI, USA

A. Bechara

University of Southern California, Los Angeles, CA, USA

G. S. Berns

Emory University, Atlanta, GA, USA

A. A. Dagher · L. Dubé

McGill University, Montreal, Canada

S. A. Huettel

Duke University, Durham, NC, USA

J. W. Kable

University of Pennsylvania, Philadelphia, PA, USA

H. Plassmann

INSEAD, Fontainebleau, France

A. Smidts

Erasmus University, Rotterdam, The Netherlands

C. Spence

University of Oxford, Oxford, UK

for different sources of heterogeneity within and across populations, suggest novel hypotheses with respect to choices and underlying mechanisms that accord with an understanding of biology, and allow for the use of neural data to make better predictions about consumer behavior. The article suggests that despite some challenges associated with incorporating neuroscience into research on consumer decision processes, the use of neuroscience paradigms will produce a deeper understanding of decision making that can lead to the development of more effective decision aids and interventions.

Keywords Consumer neuroscience · Decision neuroscience

Neuroscience has become both a useful tool and a source of theory development and testing in decision-making research. Some researchers express high hopes that a neuroscience perspective will provide a deeper understanding of marketing and consumer decision making, whereas other researchers appear more skeptical. In this paper, we tackle two related questions. How can insights from neuroscience shape future theory and models in consumer decision making? How should neuroscience methods be integrated into the research methodology of consumer decision making? We argue that neuroscience provides constraints to facilitate better theory development, provides new empirical tests of standard theoretical claims, can provide explanations for observed heterogeneity within and across populations, and can provide a mechanism for considering the physiological context and the role of constructs such as hunger, stress, and social influence on consumer choices and preferences. We present a few key arguments and examples of how the use of neuroscience paradigms can illuminate our understanding of decision processes. Our primary goal is to appeal to a broad audience and to stimulate further study in this important area. Readers, who want background on specific measures, paradigms, methods, and reviews of recent findings related to decision neuroscience, should consult Glimcher et al. (2009) or Vartanian and Mandel (2011).

1 The role of neuroscience in consumer decision making

The prospect of turning to the biological variables of neuroscience to inform models of marketing and consumer decision making may, at first, seem far-fetched. Indeed, some economists (e.g., Gul and Pesendorfer 2008) have argued that because economic and decision-making models tend to be silent about the underlying biological mechanisms, neuroscience and biological variables would be irrelevant in theory testing. We categorically disagree with this view.

Decision-making research has benefited from the revealed preferences perspective, which follows the behaviorist tradition of focusing on the observation of what people actually choose (or state that they will choose). This perspective has gone a long way in promoting empirically testable theory. In its extreme form, the revealed preference approach ignores the black box in which decisions are made. However, this view has been somewhat limiting. Many researchers build models about the processes

occurring inside the black box, but under the revealed preferences approach those models are evaluated using data from the output stage only. While some decision scientists have been reluctant to consider data in addition to choice, others, including many in marketing, consider additional variables (e.g., attitudes, memory, stated intentions, willingness to pay, response time, and priming manipulations) to be important in theory development and empirical observation. These additional variables can facilitate insights because they provide context and testable constraints. As judgment and decision-making research has demonstrated, there can be empirically testable hypotheses about the workings and mechanisms inside the black box, especially when coupled with a revealed preferences approach.

We propose that neuroscience adds value to decision-making research by enhancing the ability to make inferences beyond our usual variables and paradigms. We assert that more comprehensive theories—those making empirically testable claims about both decision processes and their output, for example, or about both biological and social variables—will be useful as the decision neuroscience field continues to develop. Two decision behaviors may be identical but may have different underlying neural circuitry. One may ask why the circuitry is relevant if at the end of the day the choice is the same. But if one understands the underlying mechanism that led to the observed choice, then one is in a better position to (a) generalize this knowledge, (b) understand contextual influences that may interact with the different neural circuitry leading to different choices, and (c) create interventions or influence those decisions more effectively. Such process knowledge can be important in many domains including policy, marketing, legal decisions, and medical decisions. In much the same way that eyetracking or verbal self-report can provide additional information about potential process, the tools of decision neuroscience can yield valuable information that can provide additional constraints on the interpretation of choice data.

Many of us have naïve intuitions about biology being fixed and unmalleable. One of the lessons from modern neuroscience is that biological variables are instead plastic and malleable. In recent years, it has become increasingly clear that although the brain is biologically based, it is also shaped by environmental, cultural, and contextual factors. Armed with knowledge of how these variables interact, decision research scholars and practitioners may be in a better position to develop more effective, more personalized and more tailored interventions and decision aids that can improve decision making.

2 Understanding heterogeneity

Understanding both inter- and intra-personal sources of heterogeneity remains a core goal of consumer behavior research, as seen in the emphasis on improving the efficiency of marketing campaigns through population segmentation (for related ideas, see Venkatraman et al. 2012). Neuroscience provides a framework in which to study and ultimately account for individual differences. Specifically, individual differences in choice need not be arbitrary or idiosyncratic. Instead, they could reflect predictable interactions between genetic markers that code for brain function (e.g., genes that shape our dopamine system), hormone and neurotransmitter levels that fluctuate with disease and state variation (e.g., sleep deprivation), and environmental

variation (e.g., stressors and life events). Many of these variables are simply outside the scope of standard models but may be highly relevant for consumer behavior.

Recent work at the interaction of biology and social sciences suggest that genes can play important roles in shaping response tendencies in dynamic interaction with the environment. Research in gene \times environment interactions and in epigenetics suggest that it is not genes alone, but the interactions between the genetic code and environment that are associated with our behaviors. For example, prior studies have found that DRD4—a dopamine receptor gene that is implicated in ADHD and other behavioral disorders—functions differently depending on the quality of parenting. Long-allele versions of the gene are associated with sensation-seeking, high-intensity pleasure seeking, and impulsivity, but only in children who were subjected to poor quality parenting (Sheese et al. 2007). There is promise in exploring genetics and gene \times environment interactions in decision-making research; examples of recent studies include Krugel et al. (2009) and Doll et al. (2011).

Functional magnetic resonance imaging (fMRI) could inform hypotheses about heterogeneity in decision processes. There could be sub-groups of individuals that approach the same decision problem using different strategies, suggesting that these different sub-groups may exhibit different patterns of brain activation. Neuroscientists have been developing techniques that discriminate brain activity in groups of subjects performing different tasks (Poldrack et al. 2009); these techniques could be used to test whether sub-groups of subjects exhibit different brain activations in the same task. This kind of approach is analogous to the market segmentation approaches familiar to consumer researchers.

Decision theorists talk much about heterogeneity in decision-making parameters (such as utility function parameters or temporal discounting parameters), and the field has developed good models to understand that heterogeneity. Neuroscience can take that understanding to a new level by adding a biological substrate to the explanation. Take for example the role of stress on decision making. Decision-making researchers may get more mileage from their paradigms if they also examine additional biological variables related to stress, e.g., hormones like cortisol. Fluctuations in cortisol can be related to decisions and so can fluctuations (possibly experimental manipulations) of the stressor. In this case, the researcher can go beyond merely examining the association of the stressor to the choice (the black box model) and say something about the underlying mechanism in the sense that the stressor initiates a biological response, which can in turn be related to the choice. For example, Mehta et al. (2010) show that fluctuations in cortisol have implications for economic decision making during negotiation and bargaining games. Stress has not typically been seen as relevant to consumer decision making. However, we believe it is an important factor in some decisions such as food choice under time or budget constraints, which may be associated with stress (e.g., grocery shopping after a stressful day at work when cortisol levels may be high). Furthermore, the cortisol variable can also act as a “manipulation check” on the experimental stressor, and could provide additional information about other processes such as habituation to the stressor. Biological data could provide evidence that the participant is habituating to the stressor, and thus provide one way to interpret an observed change in the choices the participant makes in the face of a repeated stressor. This is a type of intrapersonal heterogeneity for which biological variables can provide additional explanatory power.

A biological variable like cortisol can also elucidate why one participant did not change their choices in the presence of a stressor but another participant did. If the participant who changed her choices responded to the stressor with elevated cortisol levels but the other participant's cortisol did not fluctuate in response to the stressor, then this points to a type of heterogeneity in, say, the effect the stressor has on the participant. If the cortisol levels are the same in the two participants, then the potential explanation that each responded to the stressor differently (at least in terms of a biological response) can be ruled out. In these examples, we have focused on cortisol as the biomarker for the stress response but there could be other biological variables and measurements, such as fMRI, that could also be informative about the heterogeneity. As can be seen through these examples, the inclusion of biological information in our behavioral decision-making models may go a long way toward accounting for the observed heterogeneity, at both intrapersonal and interpersonal levels, in decision behavior.

3 Understanding value and its computation

A major question in decision neuroscience has involved understanding how goods are represented and valued by consumers. A number of studies have documented the key role of the orbitofrontal cortex in the valuation process (e.g., Kringelbach 2005; Kable and Glimcher 2009). An open question in the neuroscience of decision making has been whether the brain has implemented a system that tracks the subjective value of items for choice, how these systems are at play when consumers “miscompute” their subjective value resulting in disadvantageous decision-making outcomes such as obesity, and whether brain activity in these systems can be consciously regulated. Answers to these questions have potentially important implications for understanding the underlying neuropsychological mechanisms of consumer decision making and for designing public policy interventions.

We illustrate with an example investigating brain systems that track decision values of hungry subjects using fMRI. Based on the results of several previous studies using monkey electrophysiology and human fMRI, one can hypothesize a priori that activity in the medial orbitofrontal cortex (mOFC) would be involved in decision value (DV) computations. To test this hypothesis, Plassmann et al. (2007, 2010) used fMRI to scan hungry people's brains while they placed bids for the right to eat 50 different junk foods (e.g., chips and candy bars) in a Becker–DeGroot–Marshak auction. The participants placed bids for the right to eat a snack at the end of the experiment in 100 different bidding trials. In each trial, they were allowed to bid \$0, \$1, \$2, or \$3 for each food item.

Plassmann et al. (2010) found that right mOFC and ventromedial prefrontal cortex (VMPFC) and right dorsolateral prefrontal cortex (DLPFC) encode for DV during choice between unhealthy but appetitive food items. However, DV computations occur when choosing among appetitive and aversive items. Since dissociations between appetitive and aversive components of value signals have been shown in other domains such as anticipatory and outcome values, it is an important question as to whether appetitive and aversive DVs are computed in similar brain regions, or in separate ones. In a follow-up study, the investigators found that activity in a common area of the mOFC/VMPFC and DLPFC correlated positively with appetitive DVs and

negatively with aversive DVs. These findings suggest that the mOFC might comprise a common valuation region that encodes for both appetitive and aversive DVs (Litt et al. 2011). These results and related ones using monetary gambles and trinkets (Chib et al. 2009), or immediate and delayed rewards (Kable and Glimcher 2007, 2010) provide evidence that the brain encodes a “common currency” that allows for a shared valuation for different categories of goods (see Kable and Glimcher 2009, for a review).

4 Understanding strategic contributions to choice

The assignment of subjective values during decision making constitutes one of the fundamental computations supporting human behavior. However, different sorts of computations may be evoked under different circumstances. Individuals may have difficulty overcoming immediate hedonic concerns to achieve a long-term goal, as in the case of decision-making disorders such as obesity, addiction, and pathological gambling, suggesting that this process may sometimes be more sensitive to immediate hedonic concerns than to long-term goals and outcomes. When this happens, strategies such as cognitive reappraisal and regulation may be required in order to modulate the computations of the value system. Different decision contexts may allow individuals more or less flexibility. Some choices may demand a rapid decision and are best suited to a simplifying, heuristic approach, while other decisions may allow for a more considered, analytic choice process. Across these and many other cases, decision makers are faced with the challenge of selecting the decision *strategy* that maximizes outcomes in the face of constraints.

Recent neuroscience work has provided insight into how individuals engage cognitive strategies to shape their decision process to modify the computations involved in determining the preference for a food, and can also diminish their motivation to obtain a food through activation of regions involved in executive control and behavioral inhibition. Similarly, there may be different roles for conscious and unconscious mechanism in our decisions, and for the role of mood, emotions, and stress. These findings may have important implications for how public policy interventions are designed to fight obesity, and also how health information on food packaging could be more effectively presented. Admittedly, there is much more research needed before we fully understand the policy implications and can develop new interventions and policies based on information from neuroscience studies.

Inter- and intra-individual variation in decision strategies can result from different underlying processes. For example, a study by Venkatraman et al. (2009) investigated how people selected between multi-attribute monetary gambles whose outcomes varied from very good (e.g., win \$100) to very bad (e.g., lose \$80). Because of the complexity of the problem—which has parallels in the multi-attribute nature of many consumer decisions—different individuals approached this decision with different strategies. Some individuals generally used information about outcomes in a largely compensatory manner consistent with standard economic models, whereas others often adopted a simpler, aspiration-level rule: “Choose the gamble that maximizes my overall chance of winning.” The authors found that these inter-individual differences were well-correlated with the response of the brain’s reward system to gains and losses, with those individuals whose reward system were most sensitive to the

valence of outcomes also showing the greatest tendency toward using the simplifying rule (i.e., choosing based on valence but ignoring magnitude). Moreover, switches from one sort of strategy to the other were associated with increased activation in the dorsomedial prefrontal cortex (DMPFC), a region broadly associated with the strategic control of behavior.

Finally, the very concept of subjective value can be parsed into distinct components, each potentially associated with distinct underlying mechanisms. Decision theory has recently been exploring the role of different types of utility in decision making (e.g., experienced utility, decision utility, remembered utility, and anticipated utility). For example, recent work indicates that rewarding visual images (e.g., photographs of attractive faces) simultaneously generate two sorts of subjective value signals within the VMPFC (Smith et al. 2010): an experienced value signal associated with the attractiveness of the face being viewed, and a DV signal proportional to an individual's relative willingness to pay small amounts of money to see attractive faces. Neuroscience can provide a framework for these different definitions and components of utility that can supplement the information that emerges from traditional the revealed preference paradigm. Neuroscience can help assess whether there is more to value than what is expressed in action. Similarly, neuroscience can inform our understanding of the mechanisms of temporal discounting, and its relation to other constructs such as impulsivity and other paradigms such as delay of gratification.

5 Hypothesis generation and constraints

A subtle detail in the previous section is that neuroscience can lead to new predictions and new paradigms for dissociating processes that may not be so easy to separate using nothing more than behavioral data and traditional research designs. It is not the case that imaging data merely lead to a search for high correlations. The food example above drew on a body of prior work to generate testable hypotheses about food choice.

More generally, neuroscience provides constraints on hypotheses that can be used to account for choice data. For example, we can reject a hypothesis about decision making that implies an unrealistic biological mechanism. A model about decision making under stress that does not have fidelity with respect to our biological understanding of the stress response would not go very far in furthering our understanding or in suggesting successful interventions for responding to stress that lead to effective decision making. Similarly, a decision-making model for stress or a model for pathological decision making that accords with our understanding of biological mechanism, but does not adequately represent decision-making processes is unlikely to yield a useful model.

Neuroscience can suggest new hypotheses, whether they be about the prediction of choices or about underlying mechanisms. One such recent example is by Ho and Spence (2009) who predicted that behavioral responses by drivers to in-car warning signals will be facilitated by designing warning systems that incorporate insights about constraints of the brain. Drawing on neuroscientific findings that humans (and other animals) pay greater attention and respond more rapidly to sensory stimuli occurring in peripersonal (i.e., close to the body) than extrapersonal space, they find support for the idea that peripersonal warning signals, compared with traditional

warning signals, afford significant performance advantages. In another example, Wadhwa et al. (2008) generated novel hypotheses, based on physiological theories of “reverse alliesthesia” and neuroscience research on the dopamine system, that consumption cues that are high in incentive value (such as sampling a food or brief experiences with hedonic cues) can strengthen subsequent goal pursuit of reward-seeking behaviors (defined as a representation of an internal state associated with a desirable outcome). The authors found support for their predictions, which ran counter to common views held by marketing practitioners and health experts, that sampling a food generally leads to lower subsequent consumption. Neuroscience also brings into decision-making research variables that have not been traditional key variables in behavioral models or paradigms. For example, recall our previous examples about the examination of hunger on our choice of food or the role of stress in decision making. The use of new hypotheses, variables, and paradigms can enrich decision-making research and potentially lead to interesting new findings.

The concept of optimization, held so dearly by many decision theorists, can also be extended by taking a biological approach. Biology may provide additional concepts and other objective functions that may lead to different predictions and conceptions of optimality. So, for example, models for foraging may provide new intuitions and predictions for food decisions. Or, a series of decisions that may not appear optimal when viewed in isolation may become evident as optimal when viewed in the context of a broader system where there are additional objectives and constraints. Behaviors and decisions leading to long-term gene propagation at the expense of the immediate and direct benefit to the individual, might not appear locally optimized but make sense from an evolutionary perspective and thus might be driven by specific biological mechanisms. New insights may come into play when analyzing the biological mechanisms underlying different strategies that a decision maker may employ (not only the usual strategy of maximizing a utility integral, but other strategies such as choosing the option with maximum probability to win or minimize maximum loss).

6 Hypothesis testing

An important issue in decision and consumer neuroscience is whether neuroscience evidence can be used to test a psychological or behavioral hypothesis, or can only speak to hypotheses about neural processes. Many have argued, quite correctly, against the validity of “reverse inference” in fMRI (e.g., Poldrack 2006). Reverse inference is concluding that participants were using a particular psychological process or experiencing a particular psychological state because of the presence of activation in a certain brain region. An example would be concluding that participants felt fear because the amygdala was active during the task. While we agree that researchers should guard against reverse inferences, such thinking can play a role in generating hypotheses for further testing.

More importantly, reverse inference is not the only way that functional imaging can be used to test hypotheses about psychological processes. Functional MRI can validly speak to whether two tasks use identical psychological processes—because if they do, they should result in similar brain activation. It is the similarity/difference of the brain activation, not the anatomical location, that is crucial here. Yoon et al. (2006) used this logic to show that thinking about brands does not engage the same

psychological processes as thinking about people since the brain activity for brand judgments was quite different from that involved in person judgments. Kable and Glimcher (2010) used similar logic to argue that people do not value immediate and delayed monetary rewards using fundamentally different mechanisms since the brain networks activated by both kinds of rewards were almost identical.

7 New types of predictions

The use of neural data in models of consumer decision making holds the promise of better predictions about consumer behavior across different time scales. Knutson et al. (2007), for example, distinguished between purchased-item trials and non-purchased-item trials, and found significant differences in nucleus accumbens (NAcc) activation during preference formation, and both MPFC and insula deactivation during price processing. In line with their *a priori* hypotheses, they then estimated brain activity in these three regions of interests and entered them as covariates in a logistic regression, along with self-report measures of preference and net value, to predict subsequent purchasing decisions. The results indicated that the full model (i.e., including the neural measures) provided significantly better predictive power, albeit one offering a small advantage over a model including only self-report measures (see also Tusche et al. 2010).

Berns and colleagues demonstrated that neural responses can be used to predict purchases that are made several years later. Berns et al. (2010) conducted a study with adolescents, from October 2006 to August 2007, in which behavioral measures of preferences and neural responses were collected while participants listened to 15-s clips of songs downloaded from <http://MySpace.com>. They found that likability ratings of songs were highly correlated with activity in the caudate nucleus (an area implicated in reward and valuation). The researchers also found that the tendency among participants to change their evaluations of a song in line with its popularity (i.e., reference group's ratings) was positively correlated with activation in the anterior insula and anterior cingulate (ACC). In a subsequent study, the investigators found that the individual neural responses (in OFC and NAcc) to songs in their initial study predicted purchase decisions by the general population assessed via total number of units sold through May 2010 (Berns and Moore 2012).

8 Multiple methods are advised

Neuroscience offers a wide range of variables and paradigms. Our suggestion is that researchers use multiple methods and paradigms from multiple disciplines. Of course, within the neuroscience domain fMRI has received much attention lately because of the compelling images it produces of the functionally related blood oxygen level-dependent responses, which reflects metabolic changes associated with neuronal activity. But, there are many other techniques within the neuroscience domain that may be useful to decision researchers. Measurements of brain structure, including diffusion tensor imaging, can provide insight into differences across individuals; e.g., as when examining changes in brain regions or their connecting pathways over the lifespan. Recordings of changes in brain electrical activity, such as electroencephalography and magnetoencephalography, provide better temporal resolution at the expense of poorer spatial resolution. Newer methods based on infrared technology

show good temporal and spatial resolution (though currently limited to recording near the surface of the cortex) in a more mobile context than fMRI.

Multiple methods not only have different strengths, in terms of the biological variables they can measure, they can also permit fundamentally different inferences. None of the techniques above allow one to infer that a neural process is necessary for a decision or that it plays a causal role in generating behavior. Lesion studies, in contrast, do test necessity. One can use lesion studies and animal models of decision making to test process and mechanism. There is also a way to mimic a lesion study by temporarily disrupting a particular brain region using repetitive transcranial magnetic stimulation (rTMS) or transcranial direct current stimulation. Further, there are psychopharmacological manipulations, the multisensory nature of perception, psychophysiological variables and genetics that can be added to the decision researcher's toolbox, each providing different paradigms, variables and constructs. Of course, there are startup and transaction costs in adopting these technologies in one's research paradigm, but those costs should be judged against the potential benefit in the more pointed theory testing that emerges from considering the interplay of biological, cognitive, affective, and behavioral variables.

An example of multiple methods comes from work by Plassmann et al. (2010) who observed decision-making-related value signals in the mOFC and DLPFC. But these value signals could be an output of decision making rather than an input. If these value signals are used in decision making, then individuals with damage to this area should behave less like value maximizers. This is exactly what Fellows and Farah (2007) observed in a choice experiment with food, famous people and colors. Patients with damage to mOFC and VMPFC were more inconsistent in their choices. Plassmann et al. (2010) demonstrated that the signals in DLPFC play an important causal role in valuation. Using rTMS while participants were involved in an economic valuation task involving the consumption of real foods, they found that applying transient disruption of the DLPFC resulted in a decrease in the values assigned to the stimuli relative to a control group. The results are consistent with the possibility that the DLPFC plays a causal role in the computation of DV at the time of choice. These results show that a manipulation of brain activity encoding DV can alter behavioral preferences of consumers.

Our general point is that the field of decision making has much to gain by taking a multidisciplinary approach to its research questions. Neuroscience is a natural discipline to add to the list of disciplines that relate to the field of judgment and decision making.

9 The neurobiology of social influence

Consumer decision making hardly ever occurs in isolation. Implicitly or explicitly, consciously or unconsciously, the social context influences choice. People demonstrate various forms of herding—alignments of the thoughts or behaviors of individuals in a group (herd) without centralized coordination (Raafat et al. 2009). The conformity literature has focused on the influence of descriptive norms, which provide information about the behavior of relevant others (such as one's peers) and are distinguished from injunctive norms that specify what "ought" to be done (e.g., "do not drink and drive"). Although the phenomenon of social conformity and the

power of descriptive norms have been studied extensively in traditional behavioral research paradigms, recent evidence points to relevant neural mechanisms. For example, Klucharev et al. (2009) used fMRI to reveal that social conformity follows principles of reinforcement learning. The results indicate that a conflict with group opinion triggers a neuronal response in the NAcc and the rostral cingulate zone (RCZ), a dorsal aspect of the MPFC similar to the ‘prediction error’ signal suggested by neuroscientific models of reinforcement learning (Schultz et al. 1997). Moreover, the amplitude of this neural response predicted the magnitude of the subsequent conforming behavioral change, and the overall size of the neural signal was related to individual differences in behavioral conformity. A follow-up study by Klucharev et al. in which they down-regulated the RCZ by means of rTMS indeed reduced conforming behavior, thus providing strong evidence of the causal role of the RCZ in social influence.

As discussed already, Berns et al. (2010) demonstrated that participants changed their opinion about a music clip when receiving information about the popularity of that clip; and that tendency to change the opinion due to the perceived opinion in their reference group was positively correlated with activation in the ACC and anterior insula. Furthermore, the MPFC was also found to be central to learning about social information (advice) and for determining the extent to which it guides behavior (Behrens et al. 2008).

The hypothesis that social conformity has a basic neural mechanism generates relevant insights for consumer behavior and marketing. First, an automatic response to deviating behavior from others makes it difficult for consumers to resist such an influence. For social norm campaigns such as encouraging people to eat healthy, drive safely, or donate their organs, one may expect that providing descriptive social norm information will generate the automatic tendency to conform. Second, understanding the dopamine system as well as conflict/error processing and its pathologies may inform us about when and for whom large effects of social influence can be expected. We speculate that if aging affects the dopamine system by weakening the reward prediction signals, then the elderly may be less affected by descriptive norm information and exhibit less social conformity. Of course, one needs to be careful making inferences about neurotransmitter systems from imaging studies and, likewise, about drawing policy implications from relatively preliminary evidence. We assert that a neuroscience perspective to the problem provides novel insights and directions for testing hypotheses, and suggests new interpretations that can be tested in subsequent studies. This is exactly the kind of generative process one likes to see in research programs.

10 Consumer neuroscience concerns

Our enthusiasm for the promise of adding neuroscience approaches to the traditional study of consumer decision making is tempered by some concerns. First, we recognize that the startup costs in terms of training are exceptionally high. Researchers and graduate students need proper training. It is more than merely learning about a new variable or a new paradigm. Doing this research properly requires a new way of thinking that must be incorporated into one’s overall theoretical perspective. This will require a change to an already over-cramped graduate school curriculum. Some of the

startup cost can be absorbed through careful collaboration but there is no substitute for the behavioral researcher learning the basic paradigms, models and analysis issues in neuroscience, and likewise for the neuroscientist collaborator to learn about the basic behavioral paradigms, models and analytic issues. Furthermore, the research is relatively expensive and requires a different type of infrastructure than is common in decision-making research.

Second, we should calibrate expectations for what the field can realistically deliver, especially for practitioners. This nascent field should not promise the royal road to perfect (or better) prediction of choice. Some people can get seduced by the dramatic brain images in *Time Magazine* and the like, and think that it must be “real science,” which they assume is better than the usual behavioral work (Weisberg et al. 2008).

The fields of decision neuroscience and consumer neuroscience are academic disciplines that use a multidisciplinary and multimodal perspective to tackle its research questions. There is no magic: one cannot peek inside a decision maker’s head and predict individual’s selection of toothpaste or tomorrow’s visit to the grocery store. We must be mindful of the limits of the techniques we use. For example, fMRI methodology cannot allow definitive inferences about the neurotransmitter system in play for a particular activation. But, our general point is that what is going on inside the head as measured by various imaging and biological correlates like genes and hormones can provide new insights and new ways to test theory. This is a great opportunity for the decision-making researcher.

Where do we go next? As the field of consumer neuroscience moves into the mainstream, we need to develop publication standards, establish training centers to educate graduate students and provide additional training for faculty who want to retool. Most importantly, the field needs to tackle the exciting research questions that are now possible with the new tools in our research toolbox. Unprecedented research opportunities are now available by adopting a multidisciplinary perspective on decision making that incorporates biological approaches.

References

- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, 456, 245–U245.
- Berns, G. S., & Moore, S. (2012). A neural predictor of cultural popularity. *Journal of Consumer Psychology*, 22, 154–160.
- Berns, G. S., Capra, C. M., Moore, S., & Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *NeuroImage*, 49, 2687–2696.
- Chib, V. S., Rangel, A., Shimojo, S., & O’Doherty, J. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39), 12315–12320.
- Doll, B., Hutchison, K., & Frank, M. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *Journal of Neuroscience*, 31, 6188–6198.
- Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex*, 17(11), 2669–2674.
- Glimcher, P., Camerer, C., Fehr, E., & Poldrack, R. (2009). *Neuroeconomics: decision making and the brain*. Amsterdam: Elsevier.
- Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In A. Caplin & A. Shoter (Eds.), *The foundations of positive and normative economics*. Oxford: Oxford University Press.

- Ho, C., & Spence, C. (2009). Using peripersonal warning signals to orient a driver's gaze. *Human Factors*, 51, 539–556.
- Kable, J., & Glimcher, P. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, 10, 1625–1633.
- Kable, J., & Glimcher, P. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, 63, 733–745.
- Kable, J., & Glimcher, P. (2010). An “as soon as possible” effect in human intertemporal decision making: behavioral evidence and neural mechanisms. *Journal of Neurophysiology*, 103, 2513–2531.
- Kharchev, V., Hytonen, K., Rijpkema, M., Smidts, A., & Fernandez, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron*, 61, 140–151.
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, 53, 147–156.
- Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience*, 6, 691–702.
- Krugel, L., Biele, G., Mohr, P., Li, S.-C., & Hecker, H. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Science*, 106, 17951–17956.
- Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2011). Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex*, 21, 95–102.
- Mehta, P.H., Yap, A., & Mor, S. (2010). The biology of bargaining: dynamic hormone changes during negotiation predict economic profit. *Talk presented at the conference for the Social and Affective Neuroscience Society*, Chicago, IL.
- Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transaction. *Journal of Neuroscience*, 27(37), 9984–9988.
- Plassmann, H., O'Doherty, J., & Rangel, A. (2010). Aversive goal values are negatively encoded in the medial orbitofrontal cortex at the time of decision-making. *Journal of Neuroscience*, 30(32), 10799–10808.
- Poldrack, R. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, 10, 59–63.
- Poldrack, R., Halchenko, Y., & Hanson, S. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. *Psychological Science*, 20, 1364–1372.
- Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in Cognitive Sciences*, 13, 420–428.
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275, 1593–1599.
- Sheese, B. E., Voelker, P. M., Rothbart, M. K., & Posner, M. I. (2007). Parenting quality interacts with genetic variation in dopamine receptor D4 to influence temperament in early childhood. *Development and Psychopathology*, 19, 1039–1046.
- Smith, D., Hayden, B., Truong, T.-K., Song, A., Platt, M., & Huettel, S. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience*, 30(7), 2490–2495.
- Tusche, A., Bode, S., & Haynes, J.-D. (2010). Neural responses to unattended products predict later consumer choices. *Journal of Neuroscience*, 30(23), 8024–8031.
- Vartanian, O., & Mandel, D. (2011). *Neuroscience of decision making*. New York: Psychology Press.
- Venkatraman, V., Payne, J., Bettman, J., Luce, M., & Huettel, S. (2009). Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron*, 62, 593–602.
- Venkatraman, V., Clithoro, J. A., Fitzsimons, G. A., & Huettel, S. A. (2012). New scanner data for brand marketers: how neuroscience can help better understand differences in brand preferences. *Journal of Consumer Psychology*, 22, 143–153.
- Wadhwa, M., Shiv, B., & Nowlis, S. M. (2008). A bite to whet the reward appetite: the influence of sampling on reward-seeking behaviors. *Journal of Marketing Research*, 45, 403–413.
- Weisberg, D., Keil, F., Goodstein, J., Rawson, E., & Gray, J. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, 20, 470–477.
- Yoon, C., Gutchess, A. H., Feinberg, F., & Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *Journal of Consumer Research*, 33, 31–40.

What is ‘neuromarketing’? A discussion and agenda for future research

Nick Lee ^{*}, Amanda J. Broderick, Laura Chamberlain

Marketing Group, Aston Business School, Aston University, UK

Received 1 February 2006; received in revised form 1 March 2006; accepted 30 March 2006

Available online 12 June 2006

Abstract

Recent years have seen advances in neuroimaging to such an extent that neuroscientists are able to directly study the frequency, location, and timing of neuronal activity to an unprecedented degree. However, marketing science has remained largely unaware of such advances and their huge potential. In fact, the application of neuroimaging to market research – what has come to be called ‘neuromarketing’ – has caused considerable controversy within neuroscience circles in recent times. This paper is an attempt to widen the scope of neuromarketing beyond commercial brand and consumer behaviour applications, to include a wider conceptualisation of marketing science. Drawing from general neuroscience and neuroeconomics, neuromarketing as a field of study is defined, and some future research directions are suggested.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Neuroscience; Neuromarketing; Neuroeconomics; Marketing; Neuroimaging

1. Introduction

Recent years have seen an explosion in the abilities of neuroscientists to directly study cortical activity in terms of frequency, time, and space. The psychological and physiological sciences have been quick to apply such techniques to make startling advances in our understanding of the brain and cognition. However, most social sciences have yet to adopt neuroimaging as a standard tool or procedure for research. In particular, while economics has begun to utilise neuroimaging techniques in its research – resulting in the creation of ‘neuroeconomics’ (e.g. Braeutigam, 2005; Kenning and Plassmann, 2005; Rustichini, 2005) – marketing science has been far slower to wake up to the benefits of imaging research, despite both fields of study sharing many common concerns regarding decision making and exchange.

There are a number of possible reasons for the lack of take-up of brain imaging methodologies in marketing science. From the perspective of the marketing academic, neuroscience and cognitive psychology in general can be intimidating subjects. Furthermore, many marketing academics may see imaging techniques as simply ‘unattainable’ to them in their own

departments. However, this is generally not the case, as most business academics work within the context of a larger university with considerable facilities for brain imaging. Even if instruments such as positron emission tomography (PET), magnetoencephalography (MEG), or functional magnetic resonance imaging (fMRI) are unavailable, electroencephalography (EEG) and galvanic skin response (GSR) technology will likely be. However, the lack of knowledge of even the existence of such techniques leads to a situation where they are not considered as potential avenues of exploration.

One possible solution to this is cross-school or departmental collaboration between business and neuroscience research groups — both in terms of project design and procedure. However, from the perspective of the neuroscience researcher, there also appear to be some barriers to collaboration. In particular, while neuroeconomics appears to have raised nary a ripple of moral concern, recent opinions on ‘neuromarketing’ within the neuroscience literature have strongly questioned the ethics of applying imaging techniques to the purpose of “finding the ‘buy button in the brain’ and ...creating advertising campaigns that we will be unable to resist” (see the July 2004 Editorial of *Nature Neuroscience*, p. 683). Emotive language such as this does little to further the possibility of academic collaboration between marketing and neuroscience researchers. Furthermore, it seems such views are reasonably widely held

* Corresponding author.

E-mail address: n.j.lee@aston.ac.uk (N. Lee).

within neuroscience research groups. This is interesting, since many of the problems investigated by neuroeconomics research are virtually identical to what a marketing researcher would recognise as part of their functional domain (cf. Braeutigam, 2005; Kenning and Plassmann, 2005; e.g. Deppe et al., 2005). Yet it is marketing, not economics, which has caused such disquiet within neuroscience circles. Unfortunately, this concern (see also the February 2004 Edition in *The Lancet*, p. 71) – while containing possibly more than a grain of truth – exhibits a fundamental misunderstanding of marketing science in an academic (rather than commercial) sense.

More specifically, without entering the long and wide-ranging debate over the scope of marketing, marketing research in business schools is essentially about understanding, explaining, and predicting individual, group, and organisational behaviour relevant to markets. Such a remit encompasses a much wider range of fields than simply how to influence consumers to buy a product. The ‘buy button’ would be a finding of interest to academic marketing researchers certainly, but then so would something like the ‘love button’ to psychological scholars. Commercial interests are of course free to apply insights from academic marketing research (as they are from psychological and economic research), but this is not necessarily the remit of the marketing academic. Unfortunately, the barely concealed disdain for the idea of ‘neuromarketing’ in the neuroscience literature is clearly based on the opinion that marketing research is a commercial activity purely designed to sell products to the public, which seems to be behind the editor of *Science*, Donald Kennedy’s, concern that “brain imaging will be used in ways that infringe personal privacy to a totally unacceptable degree” (*The Lancet*, February 2004).

This note is aimed at providing a scholarly perspective on the emerging and controversial field of ‘neuromarketing’. In doing so, we aim to define what we feel neuromarketing itself is, as well as provide a brief overview of the prior work in the area. Following this, we will set out a number of key issues within marketing research which neuroimaging is likely to help provide insight into. These problems are intended to highlight how collaboration between neuroimaging and marketing researchers can advance our knowledge of many key areas pertaining not only to consumer choice, but how we interact, relate, and behave in the context of markets and organisations. Our aim is not to set boundaries on what can and cannot be investigated in such a context, but to try to encourage and inspire thought about how neuroimaging can enhance our understanding of what is – for good or ill – an unavoidable part of contemporary society.

2. Exploring and delineating the scope of neuromarketing

In recent times, ‘neuromarketing’ has come to mean the application of neuroimaging techniques to sell products, or to — as *The Lancet* puts it “dazzle potential clients with snazzy imaging technology” (February 2004, p. 71). A number of agencies have emerged offering neuroimaging (particularly fMRI) solutions to commercial marketing problems. In the US,

BrightHouse has developed a particularly high profile, while in the UK Neurosense and Neuroco have also recently opened for business. Furthermore, the Centre for Experimental Consumer Psychology at University of Wales (Bangor) collaborates with many consumer goods firms, including Unilever. Unfortunately, much of the output of such centers is commercially sensitive, meaning that there is little information available about what they actually do, even though newspapers and other outlets find such ideas compelling. Nevertheless, neuromarketing agencies have been involved in work as diverse as evaluating car preferences for DaimlerChrysler (Erk et al.’s published output of which is referred to below), the relationship between smells and colors of food products, and which advertising media are most likely to be successful in delivering different types of messages.

It is evident that the idea of evaluating the neurological correlates of consumer behaviour has caused considerable excitement within the marketing profession (e.g. Marketing Week, 2005; Mucha, 2005). Articles such as these, and the aforementioned editorials in the neuroscience literature, give the impression that neuromarketing is solely the application of neuroimaging to consumer behaviour, and how we respond to brands and advertisements. Yet even a cursory glance at the academic literature will show the scope of marketing research to be considerably broader than the response to products, brands and advertising, and even consumer behaviour in general. Any definition of neuromarketing must take into account this diversity of research. Neuroeconomics defines itself as “the application of neuroscientific methods to analyze and understand economically relevant behaviour” (Kenning and Plassmann, 2005, p. 344). Following this lead, neuromarketing as a field of study can simply be defined as the application of neuroscientific methods to analyze and understand human behaviour in relation to markets and marketing exchanges. Such a definition has two main upshots: firstly, it moves consideration of neuromarketing away from being solely the use of neuroimaging by commercial interests for their benefit; secondly, the scope of neuromarketing research is widened from solely consumer behaviour, to include many more avenues of interest, such as inter and intra-organisational research, which are common in the marketing research literature.

The contribution neuroscientific methods can make to understanding of marketing-relevant human behaviour is likely to be considerable. The advantages of physiological measurement for marketing have been noted for at least two decades (e.g. Weinstein et al., 1984). In particular, the self-assessment measures commonly used in marketing research rely totally on the ability and willingness of the respondent to accurately report their attitudes and/or prior behaviours (Petty and Cacioppo, 1983). Physiological responses, however, can be collected when respondents are directly participating in the behaviour, are difficult for subjects to control (although not difficult to affect), and although there are individual differences in physiological responding, variations in social situations and stimuli have also been shown to have a powerful effect across individuals (Cacioppo and Petty, 1985). As seen above though, neuromarketing has not been without critics and, even within academic

circles, concerns have been raised over the ability of neurological methods to adequately take into account the panoply of relevant variables in marketing theories (e.g. Stewart, 1984; 1985).

Despite its vast potential, it is clear that prior applications of neuroimaging within the marketing literature have been solely focussed on brands and consumer behaviour. In particular, EEG has been used to explore reactions to TV advertisements in a number of ways. For example, Young (2002) explored whether specific moments within ads are primarily responsible for brand development and attention. Memory and information processing have also been of interest, with Rossiter et al. (2001) using EEG to show that certain visual scenes – showing fastest activation in left frontal cortices – are also better recognised. In the neuroscience literature, Ioannides et al. (2000) and Ambler et al. (2000) report the results of MEG experiments showing how cognitive and affective advertisements elicit activity in different cortical centers. Taken together, such findings suggest that different aspects or types of advertising generate significantly different types of brain activity, possibly leading to differences in recall and/or other measures of ad effectiveness. Yet such research is piecemeal at present.

Consumer choice-making has also proved a popular subject for neuroimaging research, although it has yet to find its way into the marketing literature. Braeutigam et al. (2001, 2004) for example have explored the difference between predictable and unpredictable choices, where predictability can be related to both the frequency of prior usage of the item, and the time gap between the choice and exposure to marketing stimuli. This research suggests that different brain regions are activated according to choice predictability, with unpredictable choices eliciting activity in regions associated with silent vocalisation and judgement of rewards. Gender differences were also found. Interestingly, recent research has suggested that a variety of brain areas are associated with pleasure and rewards (e.g. Senior, 2003), and a number of these areas have been implicated in prior research. Erk et al. (2002) found that objects of high social value (sports cars) resulted in higher reward center activity (orbitofrontal cortices, anterior cingulate regions, occipital cortices) than lesser-valued objects such as small cars. Finally, in a study which received substantial attention, McClure et al. (2004) discovered that there was a higher preference for Coke over Pepsi, and also the recruitment of emotion and affect-related areas of the brain (hippocampus and dorsolateral prefrontal cortex), when respondents were told they were drinking Coke. However, blind testing suggested no such thing. Such work reinforces the complexity of choice-making, as well as the value of emotional, situational, and informational resources.

3. Some directions for scholarly neuromarketing research

Research in marketing is considerably broader than simply exploring end consumers and their decision making though. The following section is aimed at giving a flavour of the types of questions deemed important by marketing scholars, where neuroimaging techniques may prove illuminating. The impor-

tance of such areas is evidenced by their appearance in the calls for research by institutes such as the *Marketing Science Institute* and the *Institute for the Study of Business Markets*, as well as in calls for papers by numerous top-level marketing academic journals. We give special attention to non-consumer level questions in an attempt to broaden the scope of debate as to the application of neuroimaging to marketing research. Interestingly, many of these questions have been investigated in the context of neuroeconomics, yet marketing research has much to offer in such areas, among others.

3.1. Trust

Trust is an issue which has been increasing in prominence within marketing for the last decade. However, while consumer trust in brands and products is of course vital, marketing research has investigated trust on many other levels. Inter-organisational dealings such as joint ventures, strategic alliances, and business-to-business buyer/seller dyads depend on mutual trust between parties. On one hand, consumer trust in marketing claims is crucial if they are to be believed, and ultimately lead to purchase behaviour from consumers. The social utility of trust is clear when one considers that firms selling ‘fair trade’, ‘organic’, or other socially beneficial products must rely on consumer trust in their claims for success. Furthermore, in an organisational context, relationships depend on mutual trust between the parties. Without trust, opportunistic behaviour dominates interactions, negating the possibility of long-term relationships between parties and again leading to a suboptimal situation for all. Marketing research has commonly conceptualised trust as more than a simple rational economic calculation (Morgan and Hunt, 1994), and it seems likely that neuroscientific methods can provide considerable insight into the nature and development of trust.

Neuroeconomic research has begun to investigate concepts of trust beyond rationality in recent times (King-Casas et al., 2005). Neuromarketing research can also be insightful to the investigation of trust. First and foremost, it is clear that – despite the centrality of trust to marketing relationships at a number of levels – controversies over the very nature of trust still exist (e.g. Ali and Birley, 1998; Geyskens et al., 1998). Neuroimaging is likely to offer considerable insight here. Research suggests that the caudate nucleus, which is often active when learning about stimuli-response relations, is involved in experimental games requiring some kind of trust (King-Casas et al., 2005). Yet is trust a simple response to a repeated positive stimulus, or something more? More interestingly, is the trust a buyer says they have in a seller, or a consumer in a product claim, similar in terms of the nature and location of brain activity to the trust that individual says they have in a close friend or family member? In particular, measuring both the spatial and temporal characteristics of neuronal activity may be important — for example does trust in an advertising claim or new business partner require increased information processing effort and time than trust in a long-term friend? This will have important implications as to the nature of trust. Furthermore, is consumer trust in claims relating to a product similar to a

purchasing agent's trust in a contract with a supplier, and in turn is this of the same nature as the purchasing agent's trust in the individual sales executive they have negotiated with? Can trust be transferred from an organisation to a representative of that organisation? Finally, does trust evolve throughout the course of an inter-organisational relationship, or with continuing loyalty of a consumer to a single brand? Is 'trust' ever truly existent in short-term marketing relationships? Exploring and understanding such questions about the nature of trust will then lead to greater ability to explore the antecedent factors to trust, and an ability to enhance firms' ability to build trust with customers and collaborators for mutually beneficial outcomes.

3.2. Pricing

Pricing is a key tool used by organisations in the positioning of their products. Commensurate with this, much marketing research has investigated the effects of price on consumers (Bijmolt et al., 2005). Despite the amount of academic knowledge available, companies appear to use little of it when setting prices, leading to suboptimal situations for both consumers and firms. Understanding the psychology of pricing is of crucial importance if firms are to make optimal decisions and in fact has considerable utility in a broader sense. Pricing research has implications for how we understand information processing in any decision context where resources and information are scarce and costs must be weighed against benefits. Recent behavioural research for example has explored errors made by consumers when they process prices ending in 0.99 rather than a whole number — suggesting that individuals pay less attention to later numbers in a sequence (Bizer and Schindler, 2005). Other research has begun to investigate the social role of price, and how individual differences can influence how prices are perceived (Amaldoss and Jain, 2005).

At this stage however, almost all pricing research is behavioural in nature, and relies on 'assumptions' about what actually occurs when individuals process pricing information. In fact, pricing seems to lend itself almost perfectly to neuroimaging research. For example, simultaneously exploring the temporal and spatial nature of brain activity may help us understand exactly why prices such as '\$4.99' are perceived as significantly cheaper than those such as '\$5.00'. Do individuals really ignore the final two digits, or are they processed in a different manner or at a later time — for example only when detailed comparative decisions must be made? Furthermore, do time or other pressures influence the processing of prices? Also, neuroimaging looks likely to provide considerable insight into the nature of price information. Is the price of products a purely rational piece of information, or does it have emotional and/or reward-based connotations? It seems likely that the price of a basic product such as sugar is very different in nature from the price of a conspicuous product such as a Nike sports shoe, or a Porsche sports car, which should be evidenced in changes in the location of brain activity when these prices are viewed alongside their associations. Research such as this will allow us not only to understand how prices are processed, but will

afford insight into all situations where seemingly rational information is processed in decision-making situations.

3.3. Negotiation

With exchange being such a central concept in marketing, negotiations are of critical importance. For example, consumers are often in situations where they must negotiate prices or other benefits with marketing operatives — especially for big ticket items such as cars, houses, and the like. Negotiation though is an unpleasant experience for many consumers, so much so that some organisations differentiate themselves by explicitly stating 'no negotiation' (Trocchia, 2004). Inter-organisational negotiations are also a key contributor to the efficient functioning of markets, whether they be for strategic alliances, short-term collaborations, or even manufacturer–supplier negotiations.

Game theory has proven of considerable interest in economic and marketing research when examining interactions in situations where differing payoffs exist which are known to participants (e.g. Welling and Kamann, 2001). Game theoretic models have also proven useful in the evolution of neuroeconomic research (Braeutigam, 2005; Kenning and Plassmann, 2005; Rustichini, 2005). Neuroeconomic research on games can offer considerable insight into cortical activity in decision making (Rustichini, 2005). However, they tend to be focussed on competitive/cooperative behaviour (McAfee and McMillan, 1996) rather than the negotiation processes which may lead to behaviour. Unfortunately, the marketing literature currently provides little insight into the underlying processes which lie behind negotiating behaviour, and how others evaluate various negotiation strategies (Trocchia, 2004).

By contrast, neuroimaging research has already begun to investigate negotiating behaviour. Specifically, evidence suggests that emotion as well as rational cognition is a major influence on negotiating behaviour, especially when offers are considered to be unfair (Sanfey et al., 2003). In a marketing context, research such as this looks likely to help understand when and how consumers (as well as organisational agents) are likely to let their emotions override their rationality in negotiating prices or other deals. This may ultimately help consumers get a better deal and reduce those times when we look back with regret at a purchase. Other (fMRI) research has suggested that those who cooperate in an exchange are more likely to exhibit activity in the areas associated with our understanding of others' intentions (McCabe et al., 2001). Extending such research using newer multi-modal methods may further enhance our ability to understand exactly why people do or do not cooperate, even in situations where it may be optimal. For example, what situations cause us to ignore other people's benefits and solely focus on our own, or vice versa? Exploring differential brain activity in both a temporal and spatial sense may provide insight here. Furthermore, what areas or types of cortical activity are associated with risky negotiating tactics or negotiation tactics deliberately intended to harm another party? Investigating the neuronal activity underlying such suboptimal behaviours may allow us to reduce their likelihood and increase mutually beneficial outcomes to negotiation.

3.4. Marketing and society: ethics

The last decade has seen an explosion of interest in the impact of various marketing activities on society, with particular focus on ethical issues within marketing. This interest has not been restricted to marketing research, but also from disciplines like communications, sociology, politics, and not least psychology. Most obvious of these areas has been advertising's impact. Ethics in marketing is not solely concerned with the impact of advertising messages on society though. Other scholars have concerned themselves with the impact of globalisation of markets, such as fair trade and ethical production. Research has also begun to consider the idea that consumers may be harmed by a constant bombardment of marketing, with overconsumption and purchase addiction being one possible result. In sales research, much research has explored unethical selling activities and the negative outcomes of such tactics.

Neuroimaging is likely to contribute to marketing ethics in many ways of which there is space to explore here but a few. First of all, research into advertising effectiveness – which has caused so much consternation in neuroscientific circles – can contribute more than just finding the aforementioned ‘buy button’ in the brain. In fact, exploring exactly what elements of an advertisement are critical to awareness, attitudes and evaluations of products, and whether these differ for different groups, should reduce firms’ reliance on the ‘blunt instruments’ of blanket coverage, shock tactics, or sexual imagery. The application of neuroscience to marketing may form a basis for understanding how human beings create, store, recall, and relate to information such as brands in everyday life. Furthermore, it may be possible to discover whether certain aspects of advertisements and marketing activities trigger negative effects, such as overconsumption. Exploring why certain individuals become compulsive credit-users could provide outcomes of considerable social utility — are there differential locations and/or times of brain activity when a purchase is made or marketing message is viewed between those who are compulsive overpurchasers and those who maintain more appropriate levels of spending? Finally, in the sales arena, can we differentiate between the brain activity of salespeople who apply highly ethical principles to their interactions, and those who would employ less ethical action? Are less ethical individuals more likely to fixate on short-term payoffs for themselves? Neuroeconomic research has investigated altruism, suggesting that cooperation is linked to activation of reward areas (Rilling et al., 2002). However, are these same areas activated when unethical salespeople for example perform an unethical act? Investigations into such problems could in fact be amongst the most compelling within neuromarketing.

4. Concluding remarks

While neuromarketing has only recently begun to concern neuroscientists, this article has shown that neuroscientific techniques have been used on an ad-hoc basis to investigate marketing problems in an academic sense for a number of years.

Furthermore, the recent interest in neuroeconomics was shown to have considerable overlap with the domain of marketing research. We have tried to show here that the popular neuroscientific perception of neuromarketing as unethical, fundamentally flawed, and potentially harmful, should not mistakenly be applied to scholarly marketing research. Instead, we see no reason why marketing research should not be able to benefit from neuroimaging at least as much, if not more, than economics research has begun to. Indeed, the field of neuromarketing should be considered as a legitimate and important area for future research, which will allow us to more fully understand human behaviour in an extremely important context. Applying neuroimaging to marketing research problems should allow us to understand far more clearly the impact of marketing techniques, as well as gain insight into key problems concerning business relationships, answers to which have previously remained elusive.

That said, it must be stressed that neuroimaging research itself is constantly evolving, both in terms of technology as well as insights into exactly what activity and processes in various areas of the brain actually mean. For example, as technology evolves we are able to measure frequency, temporal, and spatial characteristics of brain activity more accurately and in a complimentary fashion, potentially leading to new insight into what were previously well-accepted brain functions and areas of activity. A field such as neuromarketing adds what could be called a ‘layer of theory’ on top of the actual cortical activity measure. It should not be forgotten that this layer of theory is essentially subjective and cannot directly ‘prove’ a posited relationship between marketing constructs. Nevertheless, better and more objective measurement and observation, as can be provided by neuroimaging in many cases, allows us to get closer to understanding what really happens in response to marketing stimuli, and in marketing-relevant situations.

The purpose of this article was to provide a perspective on neuromarketing which was concerned not with commercial applications, but with developing a greater understanding of a critical area of contemporary human society. While we understand the concern amongst neuroscientists regarding inappropriate application of their techniques, we looked to show that neuromarketing itself can be a valid field of study, and a rich source of problems to be investigated using insight from neuroimaging. We hoped to stimulate greater attention to neuromarketing issues within both neuroimaging and marketing research groups, as well as to expand the scope of debate and discussion on neuromarketing and other applications of neuroimaging. Both fields have much to learn from each other’s perspective, and scholarly neuromarketing research, conducted in a collaborative and non-judgemental spirit, is likely to offer us much insight into how humans behave during what is a large part of our modern lives.

References

- Ali, H., Birley, S., 1998. The role of trust in the marketing activities of entrepreneurs establishing new ventures. *J. Market. Manag.* 14, 749–763.

- Amaldoss, W., Jain, S., 2005. Pricing of conspicuous goods: a competitive analysis of social effects. *J. Mark. Res.* 42, 30–42.
- Ambler, T., Ioannides, A., Rose, S., 2000. Brands on the brain: neuro-images of advertising. *Bus. Strateg. Rev.* 11, 17–30.
- Bijmolt, T.H.A., van Heerde, H.J., Pieters, R.G.M., 2005. New empirical generalizations on the determinants of price elasticity. *J. Mark. Res.* 42, 141–156.
- Bizer, G.Y., Schindler, R.M., 2005. Direct evidence of ending-digit drop-off in price information processing. *Psychol. Mark.* 22, 771–802.
- Braeutigam, S., 2005. Neuroeconomics—from neural systems to economic behaviour. *Brain Res. Bull.* 67, 355–360.
- Braeutigam, S., Stins, J.F., Rose, S.P.R., Swithenby, S.J., Ambler, T., 2001. Magnetoencephalographic signals identify stages in real-life decision processes. *Neural Plast.* 8, 241–253.
- Braeutigam, S., Rose, S.P.R., Swithenby, S.J., Ambler, T., 2004. The distributed neuronal systems supporting choice-making in real-life situations: differences between men and women when choosing groceries detected using magnetoencephalography. *Eur. J. Neurosci.* 20, 293–302.
- Cacioppo, J.T., Petty, R.E., 1985. Physiological responses and advertising effects: is the cup half full or half empty? *Psychol. Market.* 2, 115–127.
- Deppe, M., Schwint, W., Kugel, H., Plassmann, H., Kenning, P., 2005. Non-linear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J. Neuroimaging* 15, 171–182.
- Editorial, 2004a. Brain scam? *Nat. Neurosci.* 7, 683.
- Editorial, 2004b. The Lancet neurology: neuromarketing, beyond branding. *Lancet* 3, 71.
- Editorial, 2005. Tell me what I think. *Mark. Week* 43 (April 21).
- Erk, S., Spitzer, A.P., Wunderlich, A.P., Galley, L., Walter, H., 2002. Cultural objects modulate reward circuitry. *NeuroReport* 13, 2499–2503.
- Geyskens, I., Steenkamp, J.E.M., Kumar, N., 1998. Generalisations about trust in marketing channel relationships using meta-analysis. *Int. J. Res. Mark.* 15, 223–248.
- Ioannides, A.A., Liu, L., Theofilou, D., Dammers, J., Burne, T., Ambler, T., Rose, S., 2000. Real time processing of affective and cognitive stimuli in the human brain extracted from MEG signals. *Brain Topogr.* 13, 11–19.
- Kenning, P., Plassmann, H., 2005. NeuroEconomics: an overview from an economic perspective. *Brain Res. Bull.* 67, 343–354.
- King-Casas, B., Tomlin, D., Anen, C., Cemerer, C.F., Quartz, S.R., Montague, P.R., 2005. Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.
- McAfee, R.P., McMillan, J., 1996. Competition and game theory. *J. Mark. Res.* 33, 263–268.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11832–11835.
- McClure, S.M., Li, J., Tomlin, D., Cypert, K.S., Montague, L.M., Montague, P.R., 2004. Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379–387.
- Morgan, R.M., Hunt, S.D., 1994. The commitment-trust theory of relationship marketing. *J. Market.* 58, 20–39.
- Mucha, T., 2005. This is your brain on advertising. *Business 2.0*, 35 (August).
- Petty, R.E., Cacioppo, J.T., 1983. The role of bodily responses in attitude measurement and change. In: Cacioppo, J.T., Petty, R.E. (Eds.), *Social Psychophysiology: A Sourcebook*. Guilford Press, New York.
- Rilling, J.K., Zeh, Z.K., Berns, G.S., Kilts, C.D., 2002. A neural basis for social cooperation. *Neuron* 35, 395–405.
- Rossiter, J.R., Silberstein, R.B., Harris, P.G., Nield, G.A., 2001. Brain-imaging detection of visual scene encoding in long-term memory for TV commercials. *J. Advert. Res.* 41, 13–21.
- Rustichini, A., 2005. Neuroeconomics: present and future. *Games Econ. Behav.* 52, 201–212.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.
- Senior, C., 2003. Beauty in the brain of the beholder. *Neuron* 38, 525–528.
- Stewart, D.W., 1984. Physiological measurement of advertising effect: an unfulfilled promise. *Psychol. Market.* 1, 43–48.
- Stewart, D.W., 1985. Differences between basic research and the validation of specific measures: a reply to Weinstein et al. *Psychol. Market.* 2, 41–49.
- Trocchia, P.J., 2004. Caving, role playing, and staying home: shopper coping strategies in a negotiated pricing environment. *Psychol. Mark.* 21, 823–854.
- Weinstein, S., Drozdenko, R., Weinstein, C., 1984. Brain wave analysis in advertising research. *Psychol. Market.* 1, 83–96.
- Welling, D., Kamann, D.F., 2001. Vertical cooperation in the construction industry: size does matter. *J. Supply Chain Manag.* 37, 28–34.
- Young, C., 2002. Brain waves, picture sorts®, and branding moments. *J. Advert. Res.* 42, 42–53.

SENSORY NEUROMARKETING



To understand how neuromarketing works, we need to first understand sensation. Which basic senses do we have? How do we experience the world around us? Can we use neuroscience to better understand, assess and influence consumers?

While many companies use odors, jingles, sound themes and other effectors in the attempt to affect consumers, less is known about how it actually works. How does an odor lead to a rush of memories from the past? Why does a 2 second jingle make you think of a particular brand? How do we experience the world as a coherent whole, rather than bits and pieces?

Sensory marketing also has its share of hype; companies that suggest that specific odor “personalities” work specific ways; analyze a company’s use of different senses (or lack thereof)

Does Touch Affect Taste? The Perceptual Transfer of Product Container Haptic Cues

ARADHNA KRISHNA
MAUREEN MORRIN*

We develop a conceptual framework regarding the perceptual transfer of haptic or touch-related characteristics from product containers to judgments of the products themselves. Thus, the firmness of a cup in which water is served may affect consumers' judgments of the water itself. This framework predicts that not all consumers are equally affected by such nondiagnostic haptic cues. Results from four studies show that consumers high in the autotelic need for touch (general liking for haptic input) are less affected by such nondiagnostic haptic cues compared to consumers low in the autotelic need for touch. The research has many implications for product and package design.

Does food served on a paper plate taste worse than the same food served on a china plate? Does mineral water served in a flimsy cup taste worse than the same water served in a firmer cup? Although the paper plate and flimsy cup may be less aesthetically appealing to a consumer, rationally speaking, the product containers should not affect the actual quality or taste of the products within those containers. Nevertheless, in this article we develop a conceptual framework that posits that haptic or touch-related characteristics of product containers may indeed be transferred to the products contained therein through consumer inferences and evaluations. This framework also predicts that not all consumers are equally affected by such irrelevant or non-diagnostic haptic cues from product containers. Rather, we propose that haptically oriented individuals, or those who tend to enjoy touching products, will be less affected by such cues because of their heightened awareness of and ability to correct for the potential impact of such cues.

Prior research has shown that the mere act of interpersonal touching affects behavioral compliance and reciprocation in many consumer contexts. For instance, Crusco and Wetzel

(1984) showed that being touched by a waiter in a restaurant increases tips. Hornik (1992) showed that consumers touched by a requester (to taste a new snack in a supermarket) tend to comply more than customers in a no-touch situation. In a similar vein, Hornik and Ellis (1988) showed that interpersonal touch increases shoppers' willingness to take part in mall intercept interviews.

Other research has explored how the act of touching products affects consumer response. McDaniel and Baker (1977) showed, for example, that a negative packaging attribute can sometimes lead to higher product quality evaluations. They found that potato chips in polyvinyl (vs. wax-coated) bags, which were harder to open, led consumers to believe the chips tasted better. It could be that harder-to-open bags were seen as sealing in the freshness of the chips and hence were diagnostic for product evaluation. In a study looking at touch by others, Argo, Dahl, and Morales (2006) showed that consumers react negatively if they believe products have already been touched by others. Peck and Childers (2003a, 2003b) demonstrated the importance of individual-level differences in haptic orientation or preference for product-based haptic information.

Thus, the effect of the perceived haptic properties of products would appear to have significant implications for consumer behavior. Research in this emerging area has concentrated primarily on touch versus no-touch conditions, rather than manipulating the quality of the touch-related input (McDaniel and Baker 1977). Moreover, research in this area has typically focused on contexts in which the haptic information is diagnostic for the target task—that is, when it provides objective information relevant to product judgment, such as touching a sweater to assess its thickness or texture (Peck and Childers 2003a).

In the current research we embark on a different direction

*Aradhna Krishna is the Isadore and Leon Winkelman Professor of Marketing at the Ross School of Business, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109-1234 (aradhna@umich.edu). Maureen Morrin is associate professor of marketing, Rutgers University, 227 Penn Street, Camden, NJ 08102 (mmorrin@rutgers.edu). Author names are in alphabetical order. Correspondence: Aradhna Krishna. The authors would like to thank Fengxia Fang, May Lwin, and Catherine Yeung for their help in data collection, and Julie Ruth for her many helpful suggestions on earlier drafts. Much of the work for this article was done while the first author was visiting the National University of Singapore.

John Deighton served as editor and Laura Peracchio served as associate editor for this article.

Electronically published October 17, 2007

by focusing on the effects of the *quality of nondiagnostic haptic input on product evaluation*. We propose a framework that outlines how the haptic effects will differ according to the individual's haptic orientation. To clarify, by nondiagnostic haptic cues we mean those that are not objectively relevant to the judgment task. For example, the fact that a beverage container feels flimsy to the touch should not affect the actual taste or quality of the beverage itself and thus would be considered a negatively valenced nondiagnostic haptic cue in the proposed framework.

Previous work has demonstrated the considerable effects that irrelevant or nondiagnostic information can have on consumer judgments (Broniarczyk and Gershoff 1997, 2003; Meyvis and Janiszewski 2002; Shiv, Carmon, and Ariely 2005; Simonson, Nowlis, and Simonson 1993; van Osselaer, Alba, and Manchanda 2004). Researchers to date, however, have left largely unexplored the impact of nondiagnostic cues in the domain of touch and the moderating effect of haptic orientation. Peck and Childers (2003a, 440) have noted the lack of research on the relationship between individual differences in haptic orientation and "whether this facilitates or inhibits the utilization of information available through other forms of input. It is not clear . . . whether haptic forms of information might differentially affect (either facilitate or inhibit) the use of other perceptual sense-based forms of information." The current research seeks to fill this void by suggesting how individual differences in haptic orientation moderate the use of nondiagnostic haptic cues in taste judgment.

RESEARCH ON HAPTICS AND THE NEED-FOR-TOUCH SCALE

Interest in the topic of haptics and consumers' orientation toward the sense of touch has grown in recent years. The term "haptic perception" implies "perceptual processing of inputs from multiple subsystems including those in skin, muscles, tendons and joints" (Wolfe, Kluender, and Levi 2006, 303). We focus more specifically on touch by hand. Recent research has demonstrated that individual consumers differ in terms of their haptic orientation, or need for touch (NFT), which is conceptually defined as "a preference for the extraction and utilization of information obtained through the haptic system" (Peck and Childers 2003a, 431). Haptic orientation is thus an individual difference variable that reflects a person's motivation to or preference for touch. Those with a higher NFT have been shown to be more confident in their product judgments when they are able to touch the products they evaluate, and they are more frustrated when they are not permitted to touch products (Peck and Childers 2003b).

Similar to the need-for-cognition (Cacioppo and Petty 1982) and need-to-evaluate (Jarvis and Petty 1996) scales, the need-for-touch scale does not assume a biological basis for this individual difference. Instead, the researchers propose a dual-motivation model (McClelland, Koestner, and Weinberger 1989) in which NFT is conceptualized as a construct

with two underlying dimensions: instrumental and autotelic. The instrumental dimension of NFT reflects analytic thought that is initiated by an explicit goal that drives behavior, typically a purchase goal (e.g., "the only way to make sure a product is worth buying is to actually touch it"). Thus, it is motivation driven. The autotelic dimension, on the other hand, is more preference driven and captures compulsive and affective thoughts and feelings intrinsic to an activity that are not elicited by a reference to unmet goals (e.g., "touching products can be fun" or "when browsing in stores, I like to touch lots of products"). Thus, the instrumental dimension is more relevant to specific product purchase occasions, whereas the autotelic dimension captures individuals' general liking for haptic input from products regardless of whether or not they face an immediate purchase goal.

Each of the two dimensions of NFT (instrumental and autotelic) is composed of six separate items, and Peck and Childers (2003b) suggest, based on the underlying theory, that researchers could employ either the composite NFT scale or one of the two subscales. Previous work in marketing has focused on both the instrumental (Citrin et al. 2003) and autotelic dimensions (Peck and Wiggins 2006; see the appendix) of NFT. In several studies reported by Peck and Wiggins (2006), for example, participants report their liking of communications (e.g., of a promotional pamphlet) and willingness to donate money or to become a member of an organization after reading a communication that did or did not have a haptic element attached to it (such as tree bark on a pamphlet for an arboretum). Across several studies, they found that for high (but not low) autotelics, a communication with (vs. without) a positively valenced touch element generally enhanced persuasion.

As in Peck and Wiggins (2006), the current research focuses on the autotelic dimension of NFT, as our primary interest is in understanding the effect of haptic input on consumers' mental processing, not necessarily in the context of an immediate purchase goal. In contrast to prior research, however, the focus here is on nondiagnostic haptic input, with a particular focus on the implications of negatively (vs. positively) valenced cues. Also, the nondiagnostic haptic cues in our research are a *natural part of the product-consumption experience*, such as the haptic characteristic of the cup or bottle in which water is served, which represents a comparatively subtle manipulation of haptic input compared to prior research (Peck and Wiggins 2006) in which the haptic cue was extraneous to the core consumption experience. Also, in contrast to prior research, we propose here that the impact of nondiagnostic haptic cues on consumers will be moderated by haptic orientation, such that the less haptically oriented individuals will be more likely to exhibit the impact of such cues on product evaluations. The conceptual framework that follows outlines the mental-processing differences between high and low autotelics that provide the rationale for these expectations.

CONCEPTUAL DEVELOPMENT

The conceptual framework developed here is based on two related theories: the two-stage model of cognition (Peracchio and Luna 2006; Raghbir and Krishna 1996) and the information-reduction hypothesis (Haider and Frensch 1996). The two-stage model of cognition suggests that, in arriving at product judgments, individuals will engage in a preliminary stage of automatic processing that is followed by a more deliberative, controlled processing stage. In the current context, an automatic judgment would first be formed in which the nondiagnostic haptic input (e.g., firmness of container) may affect product judgment (e.g., of the water). Then, in the second stage of processing, a more-controlled effort would be undertaken, during which the decision maker, if sufficiently knowledgeable and able, would realize that the haptic input is nondiagnostic in nature and should thus be discounted, given less weight, or entirely ignored in the product judgment.

The information-reduction hypothesis explains how high- and low-autotelic individuals would differ in terms of this two-stage process. Specifically, this theory suggests that a difference would emerge in the second, or controlled, stage of processing due to differences in prior experience. Information-reduction theory suggests that repeated practice of a task improves both speed and quality of performance, with this relationship evident across a variety of cognitive tasks (Anderson 1982, 1987; Lassaline and Logan 1993; Logan and Etherton 1994; Logan and Klapp 1991). More recently, Haider and Frensch (1996) showed that changes in the speed and quality of performance are due in part to a reduction in the amount of task information that is processed. Thus, individuals learn, with practice, to become more selective in their use of information when making judgments. Increased practice thereby enables the decision maker to better distinguish between task-relevant versus task-redundant information in this second, more-controlled stage of processing. As a result, more-practiced individuals limit their processing to only task-relevant information in the second stage.

Peck and Childers (2003a) have shown that high (vs. low) NFT individuals are more likely to touch objects, are more adept at using touch to gather information, and are likely to form richer mental product representations that include haptic properties. They suggest that such haptically oriented individuals have higher chronic accessibility to stored haptic information while using less of their cognitive-processing capacity. In addition, their richer mental representation of haptic-related information allows them to recognize more easily when haptic information is diagnostic for the task. We propose that high autotelics will process more haptic information overall than will low autotelics. However, because of their lower need of processing capacity for haptic information, and their greater ability to assess the diagnosticity of haptic input, high autotelics will discount the haptic input that is nondiagnostic to the task and focus their evaluations more on information that is diagnostic to the task. In contrast, since low autotelics are less practiced in processing haptic information, they expend greater resources to retrieve haptic-re-

levant information from memory and thus have less cognitive capacity available to focus on other information. Also, they will not recognize as easily when the haptic information is diagnostic. As such, low autotelics will be more influenced by nondiagnostic haptic information.

Put in simpler terms, people who inherently like to touch and feel objects do so very often (compared to people who do not care as much about touching objects). Over time, they develop an expertise in understanding when touch is diagnostic to a task and when it is not; that is, they know when the touch of a product is related to the inherent product quality and when it is not. Thus, for example, they would realize that the haptic qualities of a sweater are likely to affect the goodness of the sweater. As such, product evaluations of these people would be less affected by semantic connotations of haptics. Thus, while high autotelics might like touching objects overall compared to low autotelics, they are less likely to be “misled” about product quality created by differences in nondiagnostic haptic input.

Based on the preceding arguments we propose the following overarching hypothesis: *The product judgments of low, but not high, autotelics will be significantly affected by low-quality nondiagnostic haptic cues.* This overarching hypothesis is further broken into two more testable hypotheses:

- H1:** The product judgments of low (but not high) autotelics will be lower for products with lower (vs. higher) quality nondiagnostic haptic cues.
- H2:** The product judgments of low (but not high) autotelics will be significantly lower after touching (vs. not) a product with a low-quality nondiagnostic haptic cue.

Recall that we have argued that while high autotelics derive more pleasure from touching, they are simultaneously more consciously aware of haptic cues and their effects on judgment. The product judgments of low autotelics should thus be affected by low-quality nondiagnostic haptic cues, whereas those of high autotelics should not. Note that our explanation depends on the differential impact of low-quality nondiagnostic haptic cues (whereas the overall level of pleasure from touching can be greater for high autotelics). We call this the diagnosticity-based explanation.

An alternative explanation for hypotheses 1 and 2 could be offered. One could argue that high autotelics love touching products. As such, they get more pleasure from the sensations associated with touching products than low autotelics do. This might imply that they get equal pleasure from the sensations associated with touching any product (whether diagnostic or not, and whether it is of low or high haptic quality). In addition, high autotelics can also imagine the pleasurable feeling they will get from touching a product even if they can only see it (and not touch it). On the other hand, low autotelics do not get the same pleasurable feeling from touching products and are more affected by the quality

of the haptic cue even if it is nondiagnostic; thus, semantic implications of haptics—that a good product will have good haptics and a bad product will have bad haptics—will influence low autotelics. As such, there will be no effect of high- versus low-quality nondiagnostic haptic input for high autotelics, but there will be for low autotelics (hypothesis 2). Also, since high autotelics can imagine the pleasure they will get from touching an object if they can see it, there will be no difference in product judgment for high autotelics in “touch” versus “no touch” conditions. However, there will be a difference for low autotelics (hypothesis 1). Note that this alternate explanation relies on high autotelics getting equal pleasure from high- versus low-quality haptic cues. We call this the pleasure-based explanation.

We next present an overview of four studies designed to test hypotheses 1 and 2 and also test whether the diagnosticity-based or the pleasure-based explanation is more consistent with our results.

OVERVIEW OF STUDIES

We conducted four studies. The first and third studies test whether hypotheses 1 and 2 hold; that is, they test whether low—but not high—autotelic consumers tend to incorporate nondiagnostic haptic cue information into their product evaluations. The second and third studies are designed to test support (or not) for the pleasure-based alternative explanation for our hypotheses. The fourth study tests more directly our explanation. It utilizes an alternative dependent measure in the form of thought protocols to understand better the process mechanisms of our explanation and our results in earlier studies. Participants in studies 1, 2, and 4 received actual sensorial input; that is, the studies involved touching product containers. Study 3 used written descriptions of haptic qualities as opposed to sensorial input. In all studies, the experimenter was blind to the hypotheses. Studies 1 and 2 used students from a large Asian metropolitan area, whereas studies 3 and 4 used students from a midwestern town.

We chose mineral water as the target product category in all studies. Mineral water is a product category with taste differences that can be distinguished; however, the distinctions are typically subtle, such that nondiagnostic haptic cues could potentially affect consumer perceptions and evaluations. Of the four material properties that are likely to encourage product touch—namely, texture, hardness, temperature, and weight (Peck and Childers 2003a; Lederman and Klatzky 1987)—we investigated the hardness attribute. We thus chose firmness or flimsiness of the cups and bottles containing mineral water as the haptic input in all of the studies.

Pretest for Studies

We conducted a pretest among 210 undergraduates to determine the types of cups to be used in the studies that involved touching. The haptic stimuli consisted of cups that differed in firmness/flimsiness but were otherwise visually identical. They were clear plastic cups of the same size and shape. They were both transparent and had no designs,

ridges, or markings of any sort. As such, the only key difference between the cups was the degree of firmness. The participants were asked to judge the two cups on quality. Specifically, they were asked, “please evaluate the quality of the cup in which the water was served on the following scale” (the scale went from 1 = low-quality cup to 9 = high-quality cup). Since nothing varied between the cups besides the firmness, the difference in quality perception can be attributed to haptics. The test was done between subjects so that each participant evaluated either the firm or the flimsy cup. The perceived quality of the firm cup ($M = 5.78$) was significantly greater than that of the flimsy cup ($M = 4.90$; $F(1, 206) = 10.24$, $p < .005$). In the pretest, we also had participants fill out the autotelic dimension of the Peck and Childers (2003a, 2003b) NFT scale. If we do a median split of the participants on the autotelic scale to classify half of them as high autotelic and the other half as low autotelic (median = 5.60), then the results do not change; that is, both high and low autotelics find the firm cup to be of better quality than the flimsy cup (p 's $< .05$), and the interaction of autotelic (high vs. low) and cup (firm vs. flimsy) is not significant ($F(1, 206) = .39$, $p > .50$).

STUDY 1: DOES NONDIAGNOSTIC HAPTIC INPUT AFFECT JUDGMENTS ONLY OF LOW AUTOTELICS?

Study 1 consisted of a taste test of mineral water. The task for participants was to evaluate the quality of a single, unbranded water sample. In this study we manipulated haptic input by varying whether or not participants were permitted to touch the product container, a plastic cup.

All participants were served the same mineral-water product in a cup that possessed a negatively valenced nondiagnostic haptic quality (i.e., a flimsy cup). Half the participants were allowed to feel the cup in which the mineral water was served, while the other half were not. Our hypothesis was that only low autotelics would exhibit lower-quality evaluations of the water when permitted to feel the flimsy cup in which it was served. This would occur because low autotelics would lack the level of conscious awareness of the high autotelics regarding the effect of haptic cues on their evaluations and would be less likely to correct for it. Thus, low autotelics would be less able to correct for the nondiagnostic haptic-cue input during the second, controlled stage of processing.

Design

Participants were 180 undergraduate non-business students who were offered a candy bar for participation. The design was a 2 (haptic orientation; high vs. low autotelic) \times 2 (haptic input; feel, do not feel flimsy cup) full factorial, between-subjects design. As stated earlier, the autotelic dimension of NFT, which relates to touch as an end in itself and corresponds to the sensory aspect of touch (Peck and Childers 2003a), was expected to be the more relevant di-

TABLE 1
STUDY 1: PRODUCT QUALITY RATINGS

Measure	Haptic orientation	Do not feel flimsy cup	Feel flimsy cup	p-value
Product quality	Low autotelics	5.57	4.93	.017
Product quality	High autotelics	5.56	5.69	.588

mension for our purposes. A median split on the mean of the six items of the autotelic dimension of the Peck and Childers (2003a, 2003b) NFT scale was used to categorize the participants as either high or low on the autotelic dimension of haptic orientation (median = 5.50 on a nine-point scale).

Haptic input was manipulated by allowing half the participants to feel and hold the cup, with the others not permitted to do so. Only a single type of cup was used to serve the water—the flimsy cup, which received a mean rating of 4.03 on a nine-point quality scale, putting it in the negative-quality domain (i.e., providing negatively valenced haptic input). Thus, if a cup effect emerges, it should be in a negative direction in terms of the effect on evaluations.

Method

Participants, one at a time, took part in a mineral-water taste test. They were recruited at a university cafeteria and led to a room with a table. The experimenter was seated behind the table, and participants were asked to sit in front of it. The experimenter had a pitcher, which always contained the same water mixture (one can of Sprite to 6 liters of tap water) on the table, filled to the same point, before the participant entered the room. The experimenter stated, "Hello. We would like you to take part in a taste test of a new mineral water that may be introduced." The participant was then told, "This taste test requires that you taste the water through a straw. Is that all right?" All participants freely agreed to the conditions of the taste test. Each participant read a product description that described the water in terms of its vitamin and mineral content and then evaluated just one water sample.

We wanted to limit the haptic input to that which was not instrumental (i.e., diagnostic) to the judgment. Thus, we wanted to assess the effect of the haptic quality of the cup and not the taste (or feel) of the lip of the cup, which could arguably affect the perceived taste of the water. Hence, in all conditions, participants were asked to consume the water through a straw. As such, the feel of the container in the hand was the only haptic characteristic that was manipulated and that could potentially affect consumer evaluations.

Once the participant was prepared (with straw), the experimenter randomly chose a cup from behind the table (invisible to the participant), placed it on the table, poured the water into the cup, and placed the cup in front of the participant for tasting. The experimenter filled the container to the same level and put the straw into the cup for all participants.

At this point, the experimenter told the participant, "okay, please drink the water given to you in small sips through the straw." In the conditions with haptic input, participants were told, "okay, please hold the cup and drink the water given to you in small sips through the straw." The participant was then asked a set of questions about the quality of the water, knowledge of water, haptic orientation, and demographics.

Results

Immediately after tasting, participants rated the mineral water in terms of its quality (one-to-nine scale anchored at "very low/very high quality"). An ANOVA was conducted on this measure as a function of the two independent variables (haptic orientation and haptic input). We initially included gender, self-reported knowledge of mineral water (one-to-nine scale anchored at "not at all/very knowledgeable"), and age as covariates that we had measured, although only gender approached significance ($p < .065$) and was retained in the model. (Note that as is done here, all subsequent analyses include only covariates that emerged significant.) Males provided directionally higher quality ratings. There was a significant main effect of haptic orientation ($F(1, 175) = 4.24, p < .05, \eta^2 = .024$), with high autotelics providing a higher mean quality evaluation overall (M (high autotelic) = 5.63 vs. M (low autotelic) = 5.25). But this result is qualified by a significant interaction of haptic orientation and haptic input ($F(1, 175) = 4.50, p < .05, \eta^2 = .025$), which showed that the low-quality haptic input (i.e., feeling the flimsy cup) elicited more-negative evaluations only among the low autotelics (M (haptic input) = 4.93 vs. M (no haptic input) = 5.57; $F(1, 175) = 5.80, p < .05$), as predicted. The low-quality haptic input had no effect on the high autotelics' evaluations of water quality (M (haptic input) = 5.69 vs. M (no haptic input) = 5.56; $F(1, 175) = .29, p > .5$; see table 1). Since the product served to all participants was identical, this result reflects the downward effect of the negatively valenced nondiagnostic haptic cue on low autotelics' product evaluations.

Study 1 shows that touching (vs. not) a flimsy cup (i.e., a low-quality haptic input) in which water is served will affect quality perceptions for the water among low autotelics but not among high autotelics. This is consistent with hypothesis 2 that the product judgments of low (but not high) autotelics will be significantly lower after touching (vs. not) a product with a low-quality nondiagnostic haptic cue.

STUDY 2: DO HIGH AUTOTELICS OBTAIN EQUAL PLEASURE FROM TOUCHING FIRM AND FLIMSY CUPS?

Study 2 was conducted to test the alternative pleasure-based explanation that could account for our pattern of results. The first, or diagnosticity-based explanation, contends that high (vs. low) autotelics receive more pleasure from touching objects, tend to touch them more, and are more consciously aware of the potential effect of haptic cues on product judgment. As a result, they are more capable of adjusting for such cues in their product judgments when they are nondiagnostic in nature, compared to low autotelics. The product judgments of low autotelics thus should be affected by nondiagnostic haptic cues, whereas those of high autotelics should not.

The alternate pleasure-based explanation, which is tested in this study, suggests that since high autotelics enjoy touching products so much, they obtain equal pleasure from touching any product—whether diagnostic or not in terms of product evaluation, and regardless of whether it has low- or high-quality haptic cues. Conversely, low autotelics do not obtain this level of pleasure from touching products and are therefore more affected by the quality of haptic cues even though they are nondiagnostic for product evaluation. The semantic implications of haptic cues—that a good product will have good haptics and a bad product will have bad haptics—will influence low autotelics.

A critical assumption of the pleasure-based explanation is that high autotelics get equal pleasure from the sensations associated with touching firm and flimsy containers. In this experiment, we directly test this assumption. To measure pleasure from the sensations associated with touching the firm versus flimsy cups, we asked participants the question, “how much did you like the feel of the cup the water was in?” (on a scale of 1 = did not like at all to 10 = liked very much). This was deemed a more appropriate question than one worded, “how much pleasure did you get from holding the firm cup?” which may have been considered odd by participants. We also did not want to use an adapted subset of the autotelic scale since then our results could be confounded with responses on the autotelic scale.

An additional aspect of this experiment is to focus on the effect of touch alone, without any visual input. Thus, we have two sensory conditions—one in which participants are blindfolded and can only touch the cup, and one where they are not blindfolded and can both touch and see the cup. The blindfolded condition was included so that we could obtain a measure of pleasure from the pure haptic sensation for high versus low autotelics. The normal (see and touch) condition was included, since that is how participants experienced stimuli in studies 1 and 3, and also how they experience stimuli in the real world.

Design

Participants were 271 undergraduate students who completed the experiment as part of a subject pool. The design

was a 2 (haptic quality, firm or flimsy cup) \times 2 (sensory input, could touch and see the glass or could touch but not see the glass) \times 2 (haptic orientation, high vs. low autotelic) full-factorial, between-subjects design. Participants' haptic orientation was measured using the NFT scale, and we performed a median split to categorize participants as high or low autotelic based on the autotelic dimension of the scale (the median on the nine-point scale was 6.00).

Method

Participants were led into a room one at a time. In the room they were asked to sit in front of a table. The experimenter was behind the table. The table had a cup on it that was three-fourths full of water. Participants were asked to pick up the cup and feel it. For the blindfolded condition, the participants were blindfolded before they entered the room. For these participants, the experimenter stated, “This test requires that I cover your eyes with a blindfold. Is that all right?” During the test, the experimenter helped these participants find the cup. When ready, participants were told, “okay, please hold the cup and get a feel of it.”

All participants were asked a set of questions about the cup, mineral water in general, and demographics, and were also administered the NFT scale. All participants answered the questions after the cup was removed (blindfolded participants did this after the cup was taken away from sight and the blindfold was removed).

Results

Participants were asked, “how much did you like the feel of the cup the water was in?” on a scale of 1 = did not like at all to 10 = liked very much. We conducted an ANOVA on their response to this question as the dependent variable, and haptic quality (firm vs. flimsy cup), visual input (blindfolded or not), and haptic orientation (high vs. low autotelic) as the independent variables. None of the covariates were significant and were thus excluded from the analysis.

There was a significant main effect of haptic quality ($F(1, 263) = 9.11, p < .005, \eta^2 = .03$), with participants liking the feel of the cup more if it was firm ($M = 6.00$) rather than flimsy ($M = 5.15$). There was also a significant effect of haptic orientation ($F(1, 263) = 6.81, p < .01, \eta^2 = .03$; see table 2 for means), with high autotelics liking the feel of the cups more overall than low autotelics (M (high) = 5.94 vs. M (low) = 5.21). Contrary to the pleasure-based explanation, we find that both high and low autotelics like the feel of the firm cup more than that of the flimsy cup: (M (firm) = 6.40 vs. M (flimsy) = 5.48; $p < .05$) for high autotelics, and (M (firm) = 5.59 vs. M (flimsy) = 4.83; $p < .05$) for low autotelics. Thus, the crucial assumption for the alternative pleasure-based explanation is not supported.

This study shows that both high and low autotelics find the “feel” of the firm cup better than that of the flimsy cup. As such, we believe that the alternate pleasure-based explanation is less likely to justify our results.

TABLE 2
STUDY 2: MEAN LIKING OF FEEL OF CUP

Haptic orientation	Blindfold condition	Firm cup	Flimsy cup	p-value
Low autotelics	Not see	6.05	5.00	.119
Low autotelics	See	5.14	4.66	.158
	Total	5.59	4.83	.043
High autotelics	Not see	6.37	5.73	.213
High autotelics	See	6.44	5.23	.064
	Total	6.40	5.48	.027

STUDY 3: DO NONDIAGNOSTIC HAPTIC CUES ALSO EMERGE WHEN THE HAPTIC CUES ARE ONLY VERBALLY COMMUNICATED?

Study 3 tests the robustness of study 1 results by changing the stimuli from actual haptic input to a written description of haptic input. Also, we use an alternate dependent variable, namely “willingness to pay,” to enhance the generalizability of results. Further, while the haptic input in study 1 was a natural part of the product-consumption experience, it was not a part of the product itself (the water was poured into a cup and thus not in the bottle). In this study, the haptic cue is part of the product itself (we look at the haptic quality of the bottle containing the water). We see whether the pattern of study 1 results—that only low autotelics tend to be affected by nondiagnostic haptic—holds when the packaging stimuli are communicated only verbally (vs. being sensorially experienced through touch). Study 3 also tests whether the experiment results are consistent with the diagnosticity-based explanation.

Design

Participants were 277 undergraduate students who completed the experimental task as part of a participant pool. The design was a 2 (haptic orientation, high vs. low autotelic) \times 2 (haptic cue description, firm vs. flimsy bottle) full factorial between subject. A median split (median = 5.95 on a nine-point scale) on the mean of the six items of the autotelic dimension of the Peck and Childers (2003a, 2003b) NFT scale was used to categorize the participants as either high or low autotelic. The nondiagnostic haptic cue in this study consisted of how the bottle in which a new mineral water will be sold is described (as either firmer/sturdier or thinner/more flimsy).

Method

All participants read the following:

A beverage firm has developed a new bottled water which will soon be introduced in your market. Please read the description of the product below and then answer the questions that follow.

Participants then read the following product description:

This sparkling mineral water originates from a spring that flows naturally to the earth's surface. The water contains minerals your body needs such as calcium, magnesium, and potassium. It also contains just a touch of fruit juice for added flavor and vitamins. A series of independent taste tests confirmed its superiority to ordinary tap water and to some of the major competitors in the bottled water category. The tests also revealed that the bottle itself feels thinner [firmer] and more flimsy [sturdier] than most other brands.

Participants were then asked how much they would be willing to pay for this water. They were informed that a similar-sized bottle of water of other brands costs anywhere from \$1.50 to \$2.00. They were then asked how much they liked the description of the bottle in which the water would be sold and how knowledgeable they were about bottled water compared to their peers. They also filled out the autotelic dimension of the NFT scale and other demographic items. The key dependent measure in this study was price willing to pay. Because this was an open-ended response, there was considerable variation in responses, with several extreme outliers. We eliminated 14 such extreme prices (e.g., \$.00, which was more than three standard deviations from the mean), resulting in a final sample of 263 participants.

Results

We conducted an ANOVA on the price consumers were willing to pay for the new bottled water as a function of the bottle description (firm vs. flimsy), haptic orientation (high vs. low autotelic), and the covariates of age, gender, and knowledge. Since none of the covariates were significant, they were omitted from the model. The only significant effect on price willing to pay was the interaction of bottle description and haptic orientation ($F(1, 259) = 4.61, p < .05, \eta^2 = .017$). Inspection of the means showed that low autotelics were willing to pay more for the bottled mineral water in the firmer bottle (\$1.72 firm vs. \$1.57 flimsy; $p < .01$), but that high autotelics’ willingness to pay was not altered by bottle firmness (\$1.62 firm vs. \$1.64 flimsy; $p > .7$; see table 3).

TABLE 3
STUDY 3: PRICE WILLING TO PAY

Measure	Haptic orientation	Firm bottle description	Flimsy bottle description	p-value
Price willing to pay	Low autotelics	\$1.72	\$1.57	.009
Price willing to pay	High autotelics	\$1.62	\$1.64	.705

Study 3 shows that imagining drinking water from a firm (vs. flimsy) bottle of water will affect willingness to pay among low autotelics but not among high autotelics. This is consistent with study 1 results and with hypothesis 1 that the product judgments of low (but not high) autotelics will be higher for products with higher (vs. lower) quality non-diagnostic haptic cues. Together, studies 1 and 3 also support our basic hypothesis that the product judgments of low, but not high, autotelics will be significantly affected by non-diagnostic haptic cues.

STUDY 4: CAN NONDIAGNOSTIC HAPTIC CUES AFFECT THOUGHTS GENERATED ABOUT THE PRODUCT?

In study 4, we test for our diagnosticity-based explanation more directly. Thus, we try to understand better the process mechanisms at work for the effect of haptic cues on product judgment. We use a different dependent measure (thoughts about the water and the tasting context) and again manipulate haptic quality, using firm and flimsy cups. We design a scenario to induce negative thoughts for the product (water in the cup). Our theory suggests that high autotelics should realize that the product (water) and the cup are separate entities and that the cup does not affect the water quality. Therefore, they should have the same proportion of negative thoughts about the water, irrespective of the cup in which it is served. However, low autotelics should be affected by the cup such that they have a lower proportion of negative thoughts about the water when it is served in a firm cup.

In this study, we also make half the participants aware of the cup as a separate entity at the start of the experiment, whereas we do not do this for the other half. Since high autotelics are already aware of the nondiagnostic nature of the cup, making them aware of the cup as a separate entity should not affect their product evaluation (both firm and flimsy cups should still be associated with an equal number of negative thoughts about the water). We expected that making low autotelics aware of the cup as a separate entity might reduce the processing differences between high and low autotelics, alerting low autotelics to the nondiagnostic nature of the cup effect and thus allowing them to correct for it, as high autotelics do.

Design

Participants were 225 undergraduate business students who took part in the experiment to fulfill a subject-pool

requirement. The study consisted of a 2 (haptic quality: firm vs. flimsy cup) \times 2 (explicit cup awareness at the start of the experiment: yes, no) \times 2 (haptic orientation: high vs. low autotelic) full factorial. As before, we measured participants' haptic orientation using the NFT scale and classified them as high or low on the basis of a median split on the autotelic dimension of the scale. Explicit cup awareness was manipulated by informing half the participants that the firm conducting the test of mineral water was also interested in the participant's opinion of the cup in which the mineral water was served.

Method

In this study, students' desks had been individually prepared with a cup of mineral water (the same formula used in the previous studies) in either a firm or flimsy cup (the same cups used in the previous studies). Upon sitting at their seats, students opened a booklet that provided them with the following instructions:

We would like you to take part in the test of a new mineral water that an airline is thinking of serving on its flights. They will be charging for the water in flight. [The airline is also interested in your opinion of the cups they use to serve the water.] The mineral water is in front of you. Note that the cup has a straw in it. Here is what we want you to do. Please follow these instructions exactly. Please pick up the cup and then take five small sips of the water through the straw and taste it carefully.

After sipping the water, each participant was asked to "write down everything that went through your mind while drinking the water (whether or not you believe it is relevant)." Thoughts listed by participants served as the main dependent variable for the study (details provided later). They also completed the NFT scale and a self-reported knowledge of mineral water scale (one-to-nine scale anchored at "not at all/very knowledgeable"). Finally, they provided their age and gender.

Note that our informing participants that "the airline will be charging for the mineral water" was done to induce negative thoughts about the water. While it was not necessary to focus on a negatively valenced situation in this study, we did so as a design choice. In study 1, we focused on touching a flimsy cup (or not), a negatively valenced haptic cue. In studies 2 and 3 we had both flimsy and firm cups. In study 4, once again, we focused on a negatively valenced situation.

Results

Valence of Thoughts. Participants' thoughts were coded, by two independent coders, as positive (e.g., "tastes pretty good"), negative (e.g., "this is not very satisfying"), or neutral (e.g., "carbonated") in valence. The few inconsistencies in coding (<5%) were discussed and agreed upon. On average, respondents listed .35 positive thoughts, .75 negative thoughts, and 2.2 neutral thoughts. The large number of neutral thoughts is consistent with mineral water not being a high-involvement product for most consumers—that is, consumers not having strong emotional reactions to it. However, among the nonneutral thoughts, there were more negative than positive thoughts listed, suggesting that our manipulation for inducing negative thoughts worked. Since the number of thoughts varies greatly by participant (the range is one to nine), we use the number of negatively valenced as a proportion of all thoughts listed by a participant as our dependent variable.

An ANOVA on the proportion of negative thoughts as a function of haptic quality (firm vs. flimsy cup), cup awareness (yes, no), and haptic orientation (high, low autotelic) was conducted. Age was the only covariate directionally ($p < .10$) significant and thus was retained in the model. Older participants exhibited a directionally higher proportion of negative thoughts. The only significant effect was the interaction of haptic quality with haptic orientation ($F(1, 216) = 3.92, p < .05, \eta^2 = .018$). Inspection of the means showed that tasting the mineral water from a firm cup reduced the proportion of negative thoughts significantly among low autotelics ($M(\text{firm}) = .183$ vs. $M(\text{flimsy}) = .309; p < .05$) but not among high autotelics ($M(\text{firm}) = .257$ vs. $M(\text{flimsy}) = .199; p > .3$; see table 4). It is interesting to note that making participants aware of the cup as a separate entity for evaluation did not affect low autotelics' judgments about the water itself, as we expected it might. It would appear that to correct for low autotelics' reliance on nondiagnostic haptic cues, an even more salient manipulation is required.

Content of Thoughts. We also conducted a content analysis of consumer mentions of various aspects of the water and tasting task. We have proposed that high autotelics process more information about the haptic qualities of objects. As such, they should mention the cup more often in

their thought listing. This, in turn, should make high autotelics more able to correct for such influences in the evaluation task.

We coded participants' mentions of how the water tasted (e.g., it tasted sweet, fruity, or flat) and other aspects of the water (e.g., it was cold or warm, it was carbonated, or how it smelled or felt in the mouth). We also coded for mentions of the cup itself. Finally, we coded for mentions of differences from expectations and also for mentions regarding ability to judge the water. Participants' mentions of each of these items were coded by two independent coders, and the few inconsistencies in coding (<10%) were discussed and agreed upon. Note that the total number of mentions of content exceeds the total number of valenced thoughts because consumers often mentioned several aspects of the water (e.g., taste, temperature, or appearance) in a single thought (e.g., "do not like the fruity aftertaste"). While this is a single (negatively valenced) thought, it is counted as mentions of several different types (i.e., having a fruity flavor, having an aftertaste, and tasting bad).

The largest proportion of mentions was related to how the water tasted and to aspects of the water itself (71% of mentions, or 3.04 mentions on average), followed by the set of mentions concerning expectations (15%, or .63 mentions on average). The next-highest incidence was for mentions regarding the ability to judge the water (4%, or .17 mentions on average) and for mentions of the cup itself (3% of mentions, or .12 mentions on average). Remember that in the awareness condition, participants were told that the airline was also interested in their opinion of the cup. As such, we would expect both high and low autotelics to mention the cup more often when they were made aware of it. Analyses of the types of mentions showed that only those related to the cup resulted in significant results, and we limit our discussion to these.

We conducted an ANOVA on mentions of the cup as a function of haptic quality (firm, flimsy cup), cup awareness (yes, no), and haptic orientation (high, low autotelic), and on the covariates. Because none of the covariates were significant, they were dropped from the analysis. The ANOVA on the number of cup-related mentions resulted in two significant main effects: ($F(1, 217) = 4.56, p < .05, \eta^2 = .021$) for haptic orientation and ($F(1, 217) = 4.58, p < .05, \eta^2 = .021$) for cup awareness. Inspection of the means

TABLE 4
STUDY 4: MEAN PROPORTION NEGATIVE THOUGHTS

Cup awareness condition	Haptic orientation	Firm cup	Flimsy cup	p-value
Aware	Low autotelics	.171	.285	.222
Aware	High autotelics	.271	.197	.378
Not aware	Low autotelics	.198	.325	.152
Not aware	High autotelics	.242	.203	.669
Total	Low autotelics	.183	.309	.050
Total	High autotelics	.257	.199	.357

shows that high autotelics mentioned the cup about twice as many times as did low autotelics ($M = .173$ vs. $M = .069$; $p < .05$); that is, about 1 in 6 high autotelics versus 1 in 12 low autotelics mentioned the cup. Also, as might be expected, those made explicitly aware of the cup mentioned the cup more often than did those who were not made aware ($M = .173$ vs. $M = .069$, respectively; $p < .05$). None of the other effects were significant.

These results provide some additional evidence that high autotelics tend to be more consciously aware of the cup in which the water is served and thus should be better able to correct for its influence on their taste evaluations. Low autotelics, in contrast, likely should be less able to mentally correct for the influence of such factors.

Summary of Study. This study shows that high autotelics are more consciously aware of the cup in which water is served. As such, they should be more able to correct for the effect of nondiagnostic haptic cues. This provides further support for the diagnosticity-based explanation for the basic hypothesis that the product judgments of low, but not high, autotelics will be significantly affected by nondiagnostic haptic cues.

Supporters of the pleasure-based explanation could argue that since high autotelics get equal pleasure from touching the firm and flimsy cups, they have the same proportion of negative thoughts for the two. However, the results from study 2 (that high autotelics do *not* get equal pleasure from touching firm and flimsy cups) suggest that this argument is less sustainable. One could also argue that high autotelics enjoy and are thus more attentive to haptic cues, of which the cup is one. Therefore, it follows by definition that they should be more likely to mention the haptic cue (the cup). However, this would still not explain why low autotelics have more negative thoughts with a flimsy versus firm cup while high autotelics do not. Regardless, future research should explore further our diagnosticity-based explanation.

GENERAL DISCUSSION

Prior research on consumer haptics has established that those who like to touch products are more influenced by the sense of touch when they should be, that is, when the haptic input is diagnostic to the task at hand (Peck and Childers 2003a, 2003b). The results of the present research suggest, however, that haptically oriented individuals will not always be more influenced by touch-related input compared to those who are less haptically oriented. Instead, we find that those who like to touch are less influenced by touch-related input when they should indeed be less influenced, that is, in the context of nondiagnostic haptic cues. Thus, the current research implies some boundary conditions to prior research findings.

We found that the nondiagnostic haptic qualities of a product package or serving container can affect how a product is evaluated; that is, such cues can indeed have an effect on product evaluation. We discovered that this process of perceptual transfer is most likely to occur for haptically non-

expert consumers, or those for whom touching products is not particularly motivating. Studies 1, 3, and 4 show that nondiagnostic haptic cues tended to affect the perceptions and evaluations only of low autotelic consumers, that is, those with a lower need for touch.

We obtained this pattern of results using different dependent measures (product judgments, price willing to pay, thoughts), as a function of both actual haptic cue input (feeling a product container) and descriptive haptic cue input (reading about a product container), and when the haptic cue was part of the product itself (water in a bottle) versus not (water poured into a disposable cup). Low autotelics would thus appear to be more vulnerable to such irrelevant cues, as they were more likely to evaluate a mineral water more negatively when touching a flimsy disposable cup (vs. a firm cup or vs. not touching the cup), and willing to pay a higher price for bottled water described as being sold in a firm (vs. flimsy) bottle. The studies give credence to our basic hypothesis that the product judgments of low, but not high, autotelics will be significantly affected by nondiagnostic haptic cues. The results are also consistent with our two more testable hypotheses: hypothesis 1, which proposes that the product judgments of low (but not high) autotelics will be lower for products with lower (vs. higher) quality nondiagnostic haptic cues, and hypothesis 2, which suggests that the product judgments of low (but not high) autotelics will be significantly lower after touching (vs. not) a product with a low-quality nondiagnostic haptic cue.

Two alternate explanations for hypotheses 1 and 2 exist: the diagnosticity-based explanation and the pleasure-based explanation. The data obtained in studies 1 and 3 support the diagnosticity-based explanation, which suggests that high autotelics realize when haptic cues are nondiagnostic and adjust for this in their product evaluation, whereas low autotelics cannot. We propose that this holds because high autotelics, who enjoy touching products more than do low autotelics and hence have acquired greater experience with touching products, are better able to discern when touch is diagnostic for product-evaluation tasks. The alternate pleasure-based explanation suggests that high autotelics get equal pleasure from both firm and flimsy cups and hence the semantic connotations of haptic cues (that the firm cup implies a better product) do not affect product evaluation. In studies 2 and 3, we provide additional support for the diagnosticity-based explanation.

Both studies 2 and 3 indicate that high and low autotelics alike prefer the feel of firm cups to flimsy cups. Thus, it does not appear to be true that high autotelics obtain equal enjoyment from touching firm and flimsy cups. In study 4, which is designed to induce negative thoughts about the product, we show that the proportion of negative thoughts generated about the product varies by cup quality for low autotelics but not for high autotelics. Both of these studies provide additional support for our explanation, although we do not claim that they "prove" our explanation or that they "disprove" the alternate explanation.

Study 3 is based on a verbal description and hence an

"imagined touch." Our results from study 3 are consistent with those from studies 1 and 4, in which the stimuli are actually touched. Similar consistent results for real and imagined products have also been obtained by Peck and Childers (2003b), who showed pictures of the product. Beyond its theoretical contribution to sensory research, the current research has considerable managerial significance. Firms such as McDonald's, Starbucks, and Dunkin' Donuts spend millions of dollars on disposable cups and bottles each year. If such firms try to save on costs by using haptically inferior packaging, this could affect consumers' perceptions of the taste or quality of the beverages they contain. These studies also suggest that if alternate versions of a product (e.g., single malt scotch vs. a blended scotch) are served in different containers at a social function to set them apart, then consumers may (mistakenly) think that the better-quality product will be served in the haptically superior container. The studies also suggest that if lower-price generic products want to be perceived as being just as good as branded products, they may want to reconsider cutting costs via lower-quality packaging.

On the other hand, if firms make thinner containers and emphasize their environmentally friendly (less wasteful) aspects, then the thin haptic may be perceived positively and may be reflected in more-positive product evaluation (the semantic association may now be "thinner is better"). Study 2 furthermore suggests that consumers do not necessarily have to touch a product to be affected by its haptic qualities. Even a verbal description of the haptic qualities in an advertisement, for instance, can affect product evaluation.

Note that NFT is a measurable individual difference (like the need for cognition) and is relatively stable across products and situations. Still, the type of product (one in which haptic attributes are diagnostic or not) or the situation (point-of-purchase [POP] signs in a store encouraging touch) may affect the motivation of an individual to touch and also influence NFT. In this manner, one could manipulate "touch" and "NFT." Thus, in situations where haptic inputs are diagnostic, consumers should be encouraged to touch the product (e.g., through POP displays or product packaging that allows touch); this will increase their NFT and subsequently increase touching. As such, consumers would make better decisions in these scenarios. Future research should also focus on other ways to manipulate NFT to encourage touching and thus better decision making when haptics are diagnostic for the product judgment.

Where haptics are nondiagnostic, low NFTs should be made aware of this fact and thus encouraged not to rely on semantic associations of touch. In other words, where it is important for low autotelics to not be misled by nondiagnostic haptics, they should be informed that the haptics are unimportant. For instance, when choosing medicines, people should focus less on packaging and more on ingredients. Low-price medicines in cheap packaging should try and make this clear on the packet so that low autotelics make the right choice. Similarly, lower-priced brands of mass-consumed food items such as cereal, orange juice,

shampoo, and moisturizer should try and focus consumer attention away from haptics and toward the contents.

Our diagnosticity-based explanation needs further research to determine whether the differences between high and low autotelics occur because high autotelics recognize when haptic input is (or is not) diagnostic or because only high autotelics have sufficient resources to adjust for the bias. Also, in study 3, the mineral-water bottle was described as being thin and flimsy. This is clearly not realistic. Hence, the effect of verbal information needs further testing. The findings also indicate the need for much future research on haptics. Additional research is needed to study the effect of positively valenced haptic properties on consumer judgment. For example, research could explore the effect of haptic cues that might be interpreted as positive or negative based on contextual information.

In terms of future research with managerial implications, additional work is required in order to better understand the interplay between brand equity and haptic cue effects. Are haptic cues relied upon to a greater or lesser extent as a function of level of brand equity? Do lesser-known brands, for example, have greater upside potential in terms of evaluation enhancement from the perceptual transfer of haptic cues? There is also a need for a better understanding of the profile of haptically oriented consumers. What other attitudinal, demographic, or behavioral characteristics are associated with haptically oriented consumers, and do they represent an easily identifiable and thus targeted audience? There are, in addition, other dimensions of haptics worth pursuing. Besides the material quality of the object (that we have explored), how does the shape of a package affect haptic feel and thus the perceptual transfer process and product evaluation? For example, do pear-shaped bottles that are easier to hold affect evaluation of the bottle content? Similarly, does the weight of the total object, including packaging, affect evaluation of the core product—will chocolate taste better if the box is heavier, perhaps connoting creamier, denser chocolates? Study 3, which shows that product evaluation can be affected by verbal descriptions of product haptics (without experiencing the product), implies that a sensory cue related to the product can set up expectations for the product. As such, it suggests a whole new domain for future research—how can a sensory cue lead to "expectations" for another sensory cue and for the object as a whole?

The current research also indicates that perceptual transfer may occur within the context of other sensory processes. Future research can explore other domains for this phenomenon—for instance, do package smell or graphics affect product quality evaluations? And how do the sensory processes interact? For example, do bed sheets feel softer to the touch (haptics) if they come in a graphically appealing package (vision)? Our research suggests that a good understanding of the effect of nondiagnostic haptic cues is clearly important for managers for their product and packaging decisions. If the haptic feel of bottles and cups in which drinks are served and/or sold affects brand inferences, taste perceptions, and reservation price, then evidently the

choice of material for packaging and for dispensing drinks has implications for managerial profits. A dawning comprehension of the importance of this decision has spawned much recent consumer research on haptics. Our hope is to add to this growing body of research by bringing further insight to the effect of the quality of nondiagnostic haptic input and to the moderating effects of individual differences in the use of haptic input.

APPENDIX

AUTOTELIC NEED-FOR-TOUCH ITEMS (PECK AND CHILDERS 2003B)

1. When walking through stores, I cannot help touching all kinds of products.
2. Touching products can be fun.
3. When browsing in stores, it is important for me to handle all kinds of products.
4. I like to touch products even if I have no intention of buying them.
5. When browsing in stores, I like to touch lots of products.
6. I find myself touching all kinds of products in stores.

REFERENCES

- Anderson, John R. (1982), "Acquisition of Cognitive Skill," *Psychological Review*, 89, 369–406.
- (1987), "Skill Acquisition: Compilation of Weak-Method Problem Solutions," *Psychological Review*, 94, 192–219.
- Argo, Jennifer, Darren Dahl, and Andrea Morales (2006), "Consumer Contamination: How Consumers React to Products Touched by Others," *Journal of Marketing*, 70 (2), 81–94.
- Broniarczyk, Susan M. and Andrew Gershoff (1997), "Meaningless Differentiation Revisited," in *Advances in Consumer Research*, Vol. 24, ed. Merrie Brucks and Deborah J. MacInnis, Provo, UT: Association for Consumer Research, 223–28.
- (2003), "The Reciprocal Effects of Brand Equity and Trivial Attributes," *Journal of Marketing Research*, 40 (May), 161–75.
- Cacioppo, John T. and Richard E. Petty (1982), "The Need for Cognition," *Journal of Personality and Social Psychology*, 42, 116–31.
- Citrin, Alka, Donald Stem Jr., Eric Spangenberg, and Michael Clark (2003), "Consumer Need for Tactile Input: An Internet Retailing Challenge," *Journal of Business Research*, 56, 915–22.
- Crusco, April H. and Christopher G. Wetzel (1984), "The Midas Touch: The Effects of Interpersonal Touch on Restaurant Tipping," *Personality and Social Psychology Bulletin*, 10 (December), 512–17.
- Haider, Hilde and Peter A. Frensch (1996), "The Role of Information Reduction in Skill Acquisition," *Cognitive Psychology*, 30 (3), 304–37.
- Hornik, Jacob (1992), "Haptic Stimulation and Consumer Response," *Journal of Consumer Research*, 19 (December), 449–58.
- Hornik, Jacob and Shmuel Ellis (1988), "Strategies to Secure Compliance for a Mall Intercept Interview," *Public Opinion Quarterly*, 52 (4), 539–51.
- Jarvis, W. Blair G. and Richard Petty (1996), "The Need to Evaluate," *Journal of Personality and Social Psychology*, 70 (1), 172–94.
- Lassaline, Mary L. and Gordon D. Logan (1993), "Memory-Based Automaticity in the Discrimination of Visual Numerosity," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 561–81.
- Lederman, Susan J. and Roberta L. Klatzky (1987), "Hand Movements: A Window into Haptic Object Recognition," *Cognitive Psychology*, 19 (July), 342–68.
- Logan, Gordon D. and Joseph L. Etherton (1994), "What Is Learned in Automatization? The Role of Attention in Constructing an Instance," *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 1022–50.
- Logan, Gordon D. and Stuart T. Klapp (1991), "Automatizing Alphabet Arithmetic. I. Is Extended Practice Necessary to Produce Automaticity?" *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 179–95.
- McClelland, David, Richard Koestner, and Joel Weinberger (1989), "How Do Self-Attributed and Implicit Motives Differ?" *Psychological Review*, 96 (4), 690–702.
- McDaniel, Carl and R. C. Baker (1977), "Convenience Food Packaging and the Perception of Product Quality," *Journal of Marketing*, 41 (4), 57–58.
- Meyvis, Tom and Christopher Janiszewski (2002), "Consumers' Beliefs about Product Benefits: The Effect of Obviously Irrelevant Product Information," *Journal of Consumer Research*, 28, 618–35.
- Peck, Joann and Terry L. Childers (2003a), "Individual Differences in Haptic Information Processing: The 'Need for Touch' Scale," *Journal of Consumer Research*, 30 (December), 430–42.
- (2003b), "To Have and To Hold: The Influence of Haptic Information on Product Judgments," *Journal of Marketing*, 67 (April), 35–48.
- Peck, Joann and Jennifer Wiggins (2006), "It Just Feels Good: Customers' Affective Response to Touch and Its Influence on Persuasion," *Journal of Marketing*, 70 (4), 56–69.
- Peracchio, Laura A. and David Luna (2006), "The Role of Thin-Slice Judgments in Consumer Psychology," *Journal of Consumer Psychology*, 16 (1), 25–32.
- Raghubir, Priya and Aradhna Krishna (1996), "As the Crow Flies: Bias in Consumers' Map-Based Distance Judgments," *Journal of Consumer Research*, 23 (June), 26–39.
- Shiv, Baba, Ziv Carmon, and Dan Ariely (2005), "Placebo Effects of Marketing Actions: Consumers May Get What They Pay For," *Journal of Marketing Research*, 42 (4), 383–93.
- Simonson, Itamar, Stephen M. Nowlis, and Yael Simonson (1993), "The Effect of Irrelevant Preference Arguments on Consumer Choice," *Journal of Consumer Psychology*, 2 (3), 287–306.
- van Osselaer, Stijn M. J., Joseph W. Alba, and Puneet Manchanda (2004), "Irrelevant Information and Mediated Intertemporal Choice," *Journal of Consumer Psychology*, 14 (3), 257–70.
- Wolfe, Jeremy M., Keith R. Kluender, and Dennis M. Levi (2006), *Sensation and Perception*, Sunderland, MA: Sinauer.

11. J. M. D. Day, R. J. Walker, O. B. James, I. S. Puchtel, *Earth Planet. Sci. Lett.* **289**, 595 (2010).
12. R. J. Walker, *Chem. Erde* **69**, 101 (2009).
13. C.-L. Chou, *Proc. Lunar Planet. Sci. Conf.* **9**, 219 (1978).
14. W. F. Bottke, H. F. Levison, D. Nesvorný, L. Dones, *Icarus* **190**, 203 (2007).
15. By taking the ratio of Os concentration in the primitive upper mantle (~ 3.3 ppb) against the average concentration of all three chondrite groups weighted equally (~ 660 ppb) (12), we estimate that the terrestrial mass addition fraction is 0.005. Assuming the mass of Earth's mantle is 4.0×10^{24} kg, the terrestrial mass addition requirement is 2.0×10^{22} kg. A similar calculation can be made for Mars. If HSEs in the martian mantle are $0.7 \times$ the terrestrial concentration [for example, 2.4 ppb Os; consistent with the current database for shergottites (8)], and the martian mantle comprises $\sim 80\%$ of the mass of the planet (5.1×10^{23} kg), the martian mass addition requirement is $\sim 2.0 \times 10^{21}$ kg. This yields a ratio for terrestrial/martian additions of 10. For the Moon, assuming HSE abundances in the lunar mantle are ≤ 20 times lower than in the terrestrial mantle (9, 10), we obtain a maximum Os concentration of 0.16 ppb. Assuming the mass of the lunar mantle is 6.9×10^{22} kg and the same average chondritic concentration for the impactors, we obtain a lunar mass addition requirement of 1.7×10^{19} kg. This yields a ratio for terrestrial/lunar additions of 1200.
16. T. W. Dahl, D. J. Stevenson, *Earth Planet. Sci. Lett.* **295**, 177 (2010).
17. Late accretion of this magnitude would have had a modest effect on the W isotopic composition of Earth. The amount of impactor(s) mass envisioned to add the HSEs to the mantle would also add $\sim 10\%$ of the total W present in the silicate Earth (assuming that $\sim 10\%$ of the W presently resides in the silicate portion of Earth and that the late accretion impactor made up $\sim 1\%$ of the mass of the mantle). The ϵ_W value of bulk-silicate Earth, defined as a part in 10,000 deviation of $^{182}\text{W}/^{184}\text{W}$ from the terrestrial ratio, is 0. If the bulk impactor had an ϵ_W value of -2.0 (chondritic) at the time of impact, as seems likely, mass balance requires that the ϵ_W value of the silicate Earth before impact was $\sim +0.2$. If the silicate Earth and Moon had the same ϵ_W value and W concentrations at the time of the Moon-forming giant impact event (16), the Moon should have a slightly higher ϵ_W value after the proposed late accretion event. The current best estimate for the ϵ_W value for the Moon is $+0.09 \pm 0.10$ (2σ) (4). Within uncertainties, this value is ambiguously identical to both the present terrestrial value and the projected pre-late accretion impact Earth.
18. W. F. Bottke *et al.*, *Icarus* **156**, 399 (2002).
19. D. Nesvorný *et al.*, *Astrophys. J.* **713**, 816 (2010).
20. N. A. Artemieva, V. V. Shulalov, *Sol. Syst. Res.* **42**, 329 (2008).
21. We calculated an average projectile retention factor of $\sim 30\%$ for planetesimals composed of 50–50 ice/rock mixtures by splitting the difference between the data in (20) for stony and water-ice projectiles striking the Moon at 18 and 25 km/s, respectively. We exclude nearly isotropic comets from our analysis because they are extremely unlikely to hit Earth and the Moon in the interval between the Moon-forming impact and the solidification of the lunar crust.
22. I. S. Puchtel, R. J. Walker, O. B. James, D. A. Kring, *Geochim. Cosmochim. Acta* **72**, 3022 (2008).
23. A. N. Halliday, in *Meteorites, Comets and Planets: Treatise on Geochemistry*, A. M. Davis, H. D. Holland, K. K. Turekian, Ed. (Elsevier, Amsterdam, 2005), vol. 1, pp. 509–557.
24. Projectiles were randomly selected from a size distribution $dN \propto D^{-4}dD$. The range of D was set to 200 to 4000 km; this was designed to provide some flexibility in case the largest projectiles deliver only ~ 80 to 90% of their HSEs to the mantle. A successful run had an Earth/Moon mass ratio >700 with the accreted mass on Earth set to $>2.0 \times 10^{22}$ kg. The Earth/Moon impact number flux ratio was set to 20 (14, 18, 19). The sizes of the largest impactors, as well as the number of objects hitting each world within a given run were treated as variables.
25. D. C. Rubie, H. J. Melosh, J. E. Reid, C. Liebske, K. Righter, *Earth Planet. Sci. Lett.* **205**, 239 (2003).
26. D. C. Rubie, C. K. Gessmann, D. J. Frost, *Nature* **429**, 58 (2004).
27. L. E. Borg, D. S. Draper, *Meteorit. Planet. Sci.* **38**, 1713 (2003).
28. E. Asphaug, *Chem. Erde* **70**, 199 (2010).
29. V. C. Bennett, A. P. Nutman, T. M. Esat, *Geochim. Cosmochim. Acta* **66**, 2615 (2002).
30. W. D. Maier *et al.*, *Nature* **460**, 620 (2009).
31. If the proposed late accretion impactor(s) added to Earth was differentiated and contained substantial metallic iron, our model requires that the metal eventually oxidized and/or remained forever isolated from the core. This is because metal removal to the core would also lead to removal of the majority of the impactor's HSE contribution. The issue of relict metal, resulting from such an event, can be resolved via oxidation. Because the amount of metal added to the mantle by our late accretion model would comprise only $\sim 0.3\%$ of the total mass of the mantle (assuming that 30% of the mass of the impactor was metal and that the mass of the impactor was 1% of the mass of the mantle), it could have been oxidized by reaction with ferric iron in the early mantle. This would imply that the $\text{Fe}^{3+}/\Sigma\text{Fe}$ ratio (where ΣFe is total Fe) of the mantle before the impact was ~ 0.15 , rather than the present estimate of ~ 0.1 , which is well within the range of ferrous/ferric ratios possible for the early Earth, based on chondritic comparisons.
32. M. M. Marinova, O. Aharonson, E. Asphaug, *Nature* **453**, 1216 (2008).
33. A. Morbidelli, W. F. Bottke, D. Nesvorný, H. F. Levison, *Icarus* **204**, 558 (2009).
34. A. Johansen *et al.*, *Nature* **448**, 1022 (2007).
35. J. N. Cuzzi, R. C. Hogan, W. F. Bottke, *Icarus* **208**, 518 (2010).
36. See figure 6, A and B, of (33). A detailed discussion of the code and approximations used in this simulation can be found in (33).
37. W. F. Bottke, D. Nesvorný, R. E. Grimm, A. Morbidelli, D. P. O'Brien, *Nature* **439**, 821 (2006).
38. The intrinsic collision probabilities (P_i) and impact velocities (V) of stony planetesimals striking one another near 1 AU were computed from the numerical runs
- described in (37). We estimate that at ~ 1 AU, $P_i = 60 \times 10^{-18} \text{ km}^{-2} \text{ yr}^{-1}$ and $V = 11 \text{ km/s}$.
39. P. Farinella, D. R. Davis, *Icarus* **97**, 111 (1992).
40. H. F. Levison *et al.*, *Nature* **460**, 364 (2009).
41. H. V. Frey, *J. Geophys. Res. Planets* **111**, E08591 (2006).
42. K. A. Holsapple, *Annu. Rev. Earth Planet. Sci.* **21**, 333 (1993).
43. The primordial Earth was assumed to have a spin period between 6 and 8 hours. The late accretion projectiles were assumed to hit at random locations with isotropic impact angles and impact velocities between 12 and 18 km/s.
44. W. R. Ward, R. M. Canup, *Nature* **403**, 741 (2000).
45. F. M. McCubbin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11223 (2010).
46. Z. D. Sharp, C. K. Shearer, K. D. McKeegan, J. D. Barnes, Y. Q. Wang, *Science* **329**, 1050 (2010); 10.1126/science.1192606.
47. F. Robert, *Space Sci. Rev.* **106**, 87 (2003).
48. We chose ordinary and enstatite chondrites from (47) because they probably resemble the nature of planetesimals in the terrestrial planet region (6, 23). It is unclear, however, whether the measured water is consistent with the overall reduced nature of the rocks. The water may be from terrestrial weathering, though there are arguments against this (47). Regardless, the cited water quantities (0.05 to 0.2 wt %) are much lower than those in carbonaceous chondrite meteorites (1 to 10 wt %) (47). It is entirely possible that a $D = 250$ - to 300-km planetesimal striking the Moon during late accretion would have such bulk quantities of water.
49. We thank A. Barr, R. Canup, R. Citron, L. Dones, H. Levison, J. Jones, H. Nekvasil, A. Morbidelli, D. Minton, S. Mojzsis, G. J. Taylor, and D. Vokouhlický for many useful discussions and two anonymous reviewers for constructive reviews. W.F.B.'s and D.N.'s participation was supported by NASA's Lunar Science Institute through a grant to the Center for Lunar Origin and Evolution at the Southwest Research Institute in Boulder, Colorado. R.J.W.'s and J.M.D.D.'s work was supported by NASA's Lunar Science Institute through contract NNA09DB33A to the Center for Lunar Science and Exploration at the Lunar and Planetary Institute in Houston, Texas, and the NASA Astrobiology Program through grant NNG04GJ49A to the Goddard Center for Astrobiology in Greenbelt, Maryland.

24 August 2010; accepted 10 November 2010
10.1126/science.1196874

Thought for Food: Imagined Consumption Reduces Actual Consumption

Carey K. Morewedge,^{1*} Young Eun Huh,² Joachim Vosgerau²

The consumption of a food typically leads to a decrease in its subsequent intake through habituation—a decrease in one's responsiveness to the food and motivation to obtain it. We demonstrated that habituation to a food item can occur even when its consumption is merely imagined. Five experiments showed that people who repeatedly imagined eating a food (such as cheese) many times subsequently consumed less of the imagined food than did people who repeatedly imagined eating that food fewer times, imagined eating a different food (such as candy), or did not imagine eating a food. They did so because they desired to eat it less, not because they considered it less palatable. These results suggest that mental representation alone can engender habituation to a stimulus.

People believe that thinking about a desirable food or drug sensitizes one to it, increasing their hedonic response to the stimulus (1). Indeed, picturing oneself eating a delicious steak elicits an increase in salivation

and the desire to eat it (2), and imagining the sight or smell of a burning cigarette increases smokers' craving (3). The increased desirability of imagined stimuli seems to similarly affect behavior: Children have greater difficulty resisting

the impulse to eat one marshmallow immediately in order to eat two a few minutes later if they can see the marshmallow while they wait (4). Although much evidence appears to support this common intuition (4–7), its accuracy is quite puzzling, because it seems to contradict decades of research examining the overlap between direct perception and mental imagery.

Perception and mental imagery differ in their source (the senses and memory, respectively), but there is great overlap within modalities. Both engage similar neural machinery and similarly affect emotions, response tendencies, and skilled motor behavior (8–11). The thought of a spider crawling across one's leg can produce the same increases in perspiration and heart rate that would result from a spider's actual presence (5). Even the mere simulation of a motor skill can result in an improvement in its subsequent performance (9, 12). Because perception and mental imagery tend to elicit similar responses, one would expect that thinking about the consumption of a stimulus should habituate one to it.

Habituation denotes the decreased physiological and behavioral responses induced by extended or repeated exposure to a stimulus (13–15). A 10th bite of chocolate, for example, is desired less than the first bite. People habituate to a wide range of stimuli, from the brightness of a light to their income (16, 17). Habituation to food occurs too quickly for it to result from digestive feedback, so it is commonly thought to occur as a result of top-down cognitive processes (such as beliefs, memories, or expectations) or pre-ingestive sensory factors (such as texture or smell) (13, 18). Given the overlap in perception and mental representation, thinking about the consumption of a food should lead people to habituate to it.

Why then might people exhibit sensitization when thinking about a stimulus? We suggest that the sensitizing effect of imagery found in previous research may be due to the kind of imagery it has used: having participants vividly imagine a single exposure to a stimulus or associated cues. This form of imagery is more analogous to the initial exposure to a stimulus that whets the appetite and induces sensitization than to the repeated experience of a stimulus necessary to engender habituation (13). We suggest that mentally simulating an experience that is more analogous to repeated exposure (such as repeatedly imagining the consumption of units of a food) might engender habituation to the stimulus. We report five experiments testing whether repeated mental simulation of experiencing a stimulus, alone, can engender habituation. Specifically, we examined the effects of repeatedly imagining the consumption of a food on subsequent consumption of that food.

Our first experiment ($N = 51$ participants) tested whether repeatedly imagining the consumption of a food would increase or decrease the amount of that food that people would subsequently consume. All participants imagined performing 33 repetitive actions, one at a time, to hold effort constant across conditions. Controls imagined inserting 33 quarters into a laundry machine. (This served as a control task because it involves motor actions similar to those involved in eating M&M chocolate candies.) Participants in a three-repetition condition imagined inserting 30 quarters into a laundry machine and then imagined eating 3 M&M's. Participants in a 30-repetition condition imagined inserting 3 quarters into a laundry machine and then imagined eating 30 M&M's. Then all participants ate ad libitum from a bowl containing 40 g of M&M's as preparation for a "taste test" (19). The M&M's were removed when participants indicated they were finished eating. The amount of food that each participant ate was surreptitiously measured on a digital scale after the experiment. In all experiments, condition assignment was random.

A between-subjects analysis of variance (ANOVA) showed that the amount of M&M's eaten by participants was influenced by the imagination induction, $F(2, 46) = 3.61, P < 0.05$ (Table 1). Planned comparisons revealed that participants in the 30-repetition condition ate significantly fewer M&M's than did participants in the three-repetition condition and participants in the control condition, $F(1, 46) = 5.81, P < 0.05$ and $F(1, 46) = 4.50, P < 0.05$, respectively (Table 1). The amount eaten by participants in the three-repetition and control conditions did not differ significantly, $F < 1$, so the imagination induction did not sensitize participants to the food. Rather, repeatedly imagining the consumption of a food habituated participants to the food.

To ensure that the results of the previous experiment were due to imagined consumption rather than to the control task, in experiment 2

($N = 51$ participants), we orthogonally manipulated the experience that participants imagined (eating M&M's or the control task) and the number of times that experience was imagined (3 or 30 repetitions). A 2(repetitions: 3, 30) \times 2(behavior imagined: eating, control) between-subjects ANOVA yielded a significant interaction, $F(1, 47) = 4.65, P < 0.05$. Planned comparisons revealed that participants who imagined eating 30 M&M's subsequently consumed fewer M&M's than did participants who imagined eating 3 M&M's, $F(1, 47) = 4.24, P < 0.05$. Repetitions of the control task performed did not affect subsequent consumption of M&M's, $F(1, 47) = 1.05, P > 0.30$. No main effects were found, F 's < 1 (Table 1).

Experiments 3 to 5 tested whether habituation or an alternative process causes repetitive imaginary consumption of a food to reduce subsequent consumption of that food. Experiment 3 ($N = 68$ participants) tested whether one must repeatedly simulate consumption of the food or whether repeated exposure to the stimulus [priming (20)] is sufficient to reduce subsequent food intake. Participants imagined eating 3 or 30 M&M's (simulating consumption) or imagined placing 3 or 30 M&M's into a bowl (priming) before consuming the candy ad libitum, as in the previous experiments.

A 2(repetitions: 3, 30) \times 2(behavior imagined: eating M&M's, moving M&M's) between-subjects ANOVA on the amount of food consumed yielded a significant interaction, $F(1, 64) = 11.17, P < 0.05$. Planned comparisons revealed that participants who imagined eating 30 M&M's consumed significantly fewer M&M's than did participants who imagined eating 3 M&M's, $F(1, 64) = 6.10, P < 0.05$; whereas participants who imagined placing 30 M&M's into a bowl consumed significantly more M&M's than did participants who imagined placing 3 M&M's into a bowl, $F(1, 64) = 5.11, P < 0.05$. No main effects were found, F 's < 1 (Table 1). The results suggest that repetitive priming is not the process by

Table 1. Mean grams of the food consumed ad libitum after the imagery induction. Values are ± 1 SD from the mean. Means within rows that do not share the same symbol (*) or (†) differ significantly ($P < 0.05$). Each M&M weighed approximately 0.8 g; each cheese cube weighed approximately 4.5 g.

Imagery induction	Repetitions		
	0	3	30
Experiment 1 (food consumed: M&M's)			
Eating M&M's	4.08* \pm 0.33	4.18* \pm 3.26	2.21† \pm 0.48
Experiment 2 (food consumed: M&M's)			
Manipulating quarters		4.31* \pm 0.78	5.55* \pm 3.86
Eating M&M's		5.57* \pm 0.90	3.23† \pm 2.20
Experiment 3 (food consumed: M&M's)			
Moving M&M's		3.87* \pm 0.60	7.00† \pm 0.54
Eating M&M's		7.59* \pm 4.45	4.28† \pm 3.05
Experiment 4 (food consumed: cheese)			
Eating M&M's		9.47* \pm 0.98	11.15* \pm 0.82
Eating cheese cubes		11.25* \pm 0.27	6.36† \pm 0.91
After test (predicted consumption: cheese)			
Eating M&M's		12.53* \pm 2.61	13.60* \pm 2.17
Eating cheese cubes		13.41* \pm 1.81	19.21† \pm 4.01

¹Department of Social and Decision Sciences, Porter Hall 208, Carnegie Mellon University, Pittsburgh, PA, USA. ²Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA, USA.

*To whom correspondence should be addressed. E-mail: morewedge@cmu.edu

which the repetitive imaginary consumption of a food reduces subsequent food intake. Rather, repetitive priming appeared to sensitize participants to the food (21).

Habituation is stimulus-specific. Habituation to a food leads to diminished consumption of that food without much affecting the consumption of other foods (15, 18). To further test whether the effect of imaginary consumption engendered habituation or merely primed a feeling of "fullness," participants in experiment 4 ($N = 41$ participants) imagined eating the food they subsequently consumed (cheddar cheese) or imagined eating a different food (M&M's) before consuming cheddar cheese.

Participants imagined eating 3 or 30 cheddar cheese cubes or M&M's and then ate ad libitum from a bowl containing 40 g of cheddar cheese cubes. A 2(repetitions: 3, 30) \times 2(food imagined: cheddar cheese, M&M's) between-subjects ANOVA on the amount of cheese that participants consumed yielded a significant interaction, $F(1, 37) = 4.99, P < 0.05$. Planned comparisons revealed that participants who imagined eating 30 cheese cubes consumed less cheese than did participants who imagined eating 3 cheese cubes, $F(1, 37) = 5.14, P < 0.05$. In contrast, participants who imagined eating 3 M&M's or 30 M&M's did not differ significantly in the amount of cheese they consumed, $F < 1$. No main effects were found, F 's ($1, 37$) $< 1.19, P$'s > 0.28 (Table 1). The stimulus-specific effect of imagined consumption provides further evidence that habituation is the process by which repetitive imaginary consumption of a food leads to a reduction in its subsequent intake.

To ensure that the results of experiments 1 to 4 were not due to experimental demand, we described the design of experiment 4 to a new sample of participants ($N = 80$), who predicted average cheese consumption in each of its four conditions. Predictors correctly anticipated that

the imagined consumption of M&M's did not influence the actual subsequent consumption of cheese, $t(79) = 0.66, P = 0.51$, but incorrectly predicted that participants who imagined eating 30 cubes of cheese would consume more cheese than would participants who imagined eating 3 cubes of cheese, $t(79) = 3.09, P < 0.01$; within-subjects ANOVA yielded a significant interaction, $F(1, 79) = 9.23, P < 0.01$ (Table 1).

Two psychological processes with distinct neural substrates appear to regulate food selection and intake. One process entails the hedonic responses to the food (liking or palatability) and may diminish intake through sensory-specific satiation (18). The other entails the motivation and appetitive drive to obtain it (wanting) and diminishes intake through habituation (15, 22, 23). In experiment 5 ($N = 81$ participants), we used a standard procedure to test changes in liking and wanting (23–25) to identify the process by which imagined consumption reduces food intake. Participants rated their liking for the imagined food before and after an imagination induction and played a reinforcement game, a measure of habituation (13, 15, 23–25). (Note that this method tested whether repeatedly imagining the consumption of a food leads one to feel disgusted by the food, which could also diminish consumption.)

First, all participants rated how much they liked cheddar cheese on a seven-point scale with endpoints: dislike extremely (1) and like extremely (7). Then, they imagined 33 repetitive tasks. Participants in a three-repetition condition imagined performing 30 repetitions of the control task (as in experiment 1) and then imagined eating 3 cheddar cheese cubes. Participants in a 30-repetition condition imagined performing three repetitions of the control task and then imagined eating 30 cheddar cheese cubes.

Then, all participants played the reinforcement game: Participants were shown a picture

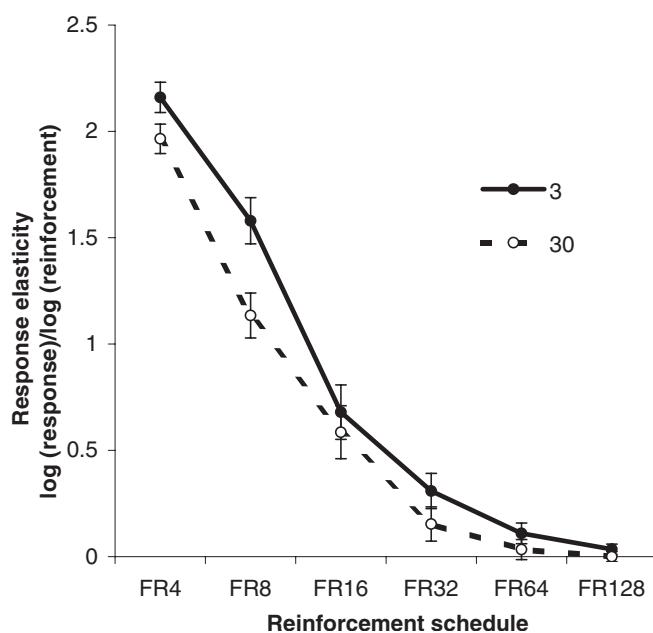
of a cheddar cheese cube and a STOP sign; they could collect points by clicking on the cheddar cheese cube. The game began with a fixed-ratio reinforcement schedule of 4 (FR4), in which every fourth click earned a point. Every time participants earned five points, the reinforcement ratio was doubled, progressing through FR8, FR16, FR32, FR64, and FR128. Participants could end the game at any time by clicking the stop sign. For each three points earned in the game, participants received one cheddar cheese cube at the end of the experiment. At the end of the game, participants re-rated how much they liked cheddar cheese on a scale identical to the scale used in the beginning of the experiment.

We calculated the difference between liking ratings before and after the imagination task to create an index of change in liking, and log-transformed responses in the reinforcement game to create an index of wanting. A 2(repetitions: 3, 30) \times 2(determinant: wanting, liking) mixed ANOVA with repetitions imagined as a between-subject factor and determinant as a within-subjects factor yielded a significant main effect for determinant $F(1, 66) = 637.10, P < 0.01$, and a marginally significant main effect for repetition, $F(1, 66) = 3.00, P < 0.1$, which were qualified by a significant interaction, $F(1, 66) = 4.82, P < 0.05$. Planned comparisons revealed that participants who imagined eating 30 cheese cubes clicked fewer times for cheese cubes in the reinforcement game ($M_{\log \text{ response}} = 3.68 \pm 1.40$) than did participants who imagined eating 3 cheese cubes ($M_{\log \text{ response}} = 4.35 \pm 1.02$), $F(1, 66) = 5.01, P < 0.05$. No difference in the change of liking was observed between the two conditions ($M_{\Delta \text{ three-repetitions}} = 0.00 \pm 0.87; M_{\Delta \text{ thirty-repetitions}} = 0.03 \pm 0.51$), $F < 1$.

Response elasticities [responses per unit price (26)] were then computed as [$\log(\text{number of responses emitted during a specific reinforcement schedule}) / \log(\text{number of reinforced responses during that reinforcement schedule})$] and are displayed in Fig. 1. A 2(Repetitions: 3, 30) \times 6(Reinforcement schedule: FR4–FR128) mixed ANOVA with repetitions as a between-subject factor and reinforcement schedules as a within-subject factor revealed that response elasticities declined linearly as reinforcement schedules progressed, $F(5, 62) = 315.08, P < 0.001$. As revealed in our previous analysis, participants who imagined eating 30 cheese cubes were less motivated to earn points toward cheese cubes than were participants who imagined eating 3 cheese cubes, $F(1, 66) = 4.69, P < 0.05$. There was a marginally significant repetitions \times reinforcement schedule interaction, $F(5, 62) = 2.31, P < 0.1$. These results suggest that imagined consumption of a food decreased its subsequent consumption through habituation because it diminished the degree to which people wanted the food; imagined consumption did not appear to affect how much they liked it.

Together, the results show that repeatedly imagining the consumption of a food leads people to habituate to it. Participants who imagined con-

Fig. 1. Motivation to earn points for cheese (expressed as response elasticity) as a function of the number of cubes of cheese (3 versus 30) that participants had imagined eating and fixed-ratio reinforcement (for example, in FR4, every fourth response was reinforced). Error bars represent ± 1 SEM.



suming more of a food were subsequently less motivated to obtain it than were participants who imagined consuming less of the food. The influence of the imagery induction on food consumption was stimulus-specific: It reduced the consumption of the food that participants imagined eating but did not affect the consumption of other foods. Finally, repetitively imagining the consumption of the food reduced wanting (the appetitive or motivational drive) for the food. Participants who imagined consuming more of the food subsequently expended less effort to obtain it.

These findings have important implications for three fields of research. First, the results suggest that mental imagery alone can engender habituation to a stimulus. In addition to its theoretical importance, this finding may allow for the development of more effective interventions to reduce cravings for unhealthy food and drugs and diminish phobic reactions to fear-inducing stimuli. Second, it has long been debated whether pre-ingestive sensory stimulation is required for habituation to occur (13, 27–29). Previous experiments in this vein have suffered from the confound of simultaneously enacting top-down processes and pre-ingestive sensory stimulation (participants ate or smelled the food to which they habituated). Mental imagery is considered a top-down cognitive process (10, 30). Thus, the results show that top-down processes can enact habituation in the absence of pre-ingestive sensory stimulation. Finally, the results show that

repeatedly simulating an action can trigger its behavioral consequences (2). Rather than increase the likelihood of enacting the simulated behavior (eating), simulation evoked the consequences of the behavior (habituation). The difference between actual experience and mental representations of experience may be smaller than previously assumed (10).

References and Notes

1. B. Soetens, C. Braet, P. Dejonckheere, A. Roets, *J. Health Psychol.* **11**, 655 (2006).
2. M. R. Dadds, D. H. Bovbjerg, W. H. Redd, T. R. Cutmore, *Psychol. Bull.* **122**, 89 (1997).
3. S. T. Tiffany, D. J. Drobis, *Addict. Behav.* **15**, 531 (1990).
4. W. Mischel, Y. Shoda, M. I. Rodriguez, *Science* **244**, 933 (1989).
5. P. Lang, *Behav. Ther.* **8**, 862 (1977).
6. D. J. Kavanagh, J. Andrade, J. May, *Psychol. Rev.* **112**, 446 (2005).
7. E. Kemps, M. Tiggemann, *J. Exp. Psychol. Appl.* **13**, 95 (2007).
8. J. Decety, J. Grèzes, *Brain Res.* **1079**, 4 (2006).
9. J. Driskell, C. Copper, A. Moran, *J. Appl. Psychol.* **79**, 481 (1994).
10. S. M. Kosslyn, G. Ganis, W. L. Thompson, *Nat. Rev. Neurosci.* **2**, 635 (2001).
11. L. E. Williams, J. A. Bargh, *Psychol. Sci.* **19**, 302 (2008).
12. E. L. Wohldmann, A. F. Healy, L. E. Bourne Jr., *J. Exp. Psychol. Learn. Mem. Cogn.* **33**, 254 (2007).
13. F. McSweeney, S. Swindell, *Psychol. Bull.* **125**, 437 (1999).
14. F. K. McSweeney, E. S. Murphy, *J. Exp. Anal. Behav.* **74**, 347 (2000).
15. L. H. Epstein, J. L. Temple, J. N. Roemmich, M. E. Bouton, *Psychol. Rev.* **116**, 384 (2009).
16. P. Brickman, D. Coates, R. Janoff-Bulman, *J. Pers. Soc. Psychol.* **36**, 917 (1978).
17. S. Frederick, G. Loewenstein, in *Well-Being: The Foundations of Hedonic Psychology*, D. Kahneman, Eds. (Russell Sage, New York, 1999), pp. 302–329.
18. B. J. Rolls, E. T. Rolls, K. Sweeney, *Physiol. Behav.* **27**, 137 (1981).
19. Materials and methods are available as supporting material on Science Online.
20. D. E. Huber, T. F. Clark, T. Curran, P. Winkielman, *J. Exp. Psychol. Learn. Mem. Cogn.* **34**, 1305 (2008).
21. R. Zajonc, *Am. Psychol.* **35**, 151 (1980).
22. K. C. Berridge, *Neurosci. Biobehav. Rev.* **20**, 1 (1996).
23. L. H. Epstein et al., *Appetite* **41**, 283 (2003).
24. R. C. Havermans, T. Janssen, J. C. Giesen, A. Roefs, A. Jansen, *Appetite* **52**, 222 (2009).
25. M. Myers Ernst, L. H. Epstein, *Appetite* **38**, 224 (2002).
26. R. J. DeGrandpre, W. K. Bickel, J. R. Hughes, M. P. Layng, G. Badger, *J. Exp. Anal. Behav.* **60**, 641 (1993).
27. S. Higgs, A. C. Williamson, P. Rotstein, G. W. Humphreys, *Psychol. Sci.* **19**, 623 (2008).
28. P. Rozin, S. Dow, M. Moscovitch, S. Rajaram, *Psychol. Sci.* **9**, 392 (1998).
29. B. J. Rolls, E. A. Rowe, E. T. Rolls, *Physiol. Behav.* **29**, 409 (1982).
30. F. Crick, C. Koch, *Nature* **375**, 121 (1995).
31. This work benefited from financial support from a grant to C.K.M. from the Berkman Faculty Development Fund at Carnegie Mellon University. We thank A. Kim, K. Khanna, M. Fei, L. Mosca, M. Christian, and Center for Behavioral Decision Research summer interns in 2008 and 2009 for help with the data collection. We thank C. Janiszewski, D. Kahneman, G. Loewenstein, B. Rolls, and L. Williams for helpful comments.

Supporting Online Material

www.sciencemag.org/cgi/content/full/330/6010/1530/DC1
SOM Text
References

27 July 2010; accepted 28 October 2010
10.1126/science.1195701

Sex Determination in the Social Amoeba *Dictyostelium discoideum*

Gareth Bloomfield,^{1,2*} Jason Skelton,² Alasdair Ivens,² Yoshimasa Tanaka,^{1,3} Robert R. Kay¹

The genetics of sex determination remain mysterious in many organisms, including some that are otherwise well studied. Here we report the discovery and analysis of the mating-type locus of the model organism *Dictyostelium discoideum*. Three forms of a single genetic locus specify this species' three mating types: two versions of the locus are entirely different in sequence, and the third resembles a composite of the other two. Single, unrelated genes are sufficient to determine two of the mating types, whereas homologs of both these genes are required in the composite type. The key genes encode polypeptides that possess no recognizable similarity to established protein families. Sex determination in the social amoebae thus appears to use regulators that are unrelated to any others currently known.

Most eukaryotes are sexual, but little is known in molecular detail about sex across most branches of the eukaryotic tree. One aspect, the genetic basis of sex determination, is well understood in several animal,

fungi, and plant lineages (1–5), but across the protozoan kingdoms we know little, and nothing in comparable detail. The social amoebae are members of the Amoebozoa and have an unusual sexual cycle that leads to the formation of dormant walled macrocysts (6) (Fig. 1, A and B). To produce a macrocyst, a pair of haploid amoebae of different sexes fuse (7) to form a diploid zygote, which then attracts surrounding haploid cells (8). These help to lay down external layers of cellulose around the developing mass of cells before being cannibalized by the zygote (9). After a period of dormancy,

the cyst germinates, releasing haploid progeny that arise most likely after meiosis and multiple mitoses (10). The population genetics of wild isolates indicate that mating and recombination are probably frequent in the wild (11).

The most-studied species of social amoeba, *Dictyostelium discoideum*, is notable for having three sexes [hereafter called mating types I, II, and III [supporting online material (SOM) text S1], as well as uncommon self-fertile homothallic strains (12–14)]. Each of the three sexes can pair with each of the other two but not with itself, giving three possible classes of zygote: type I/type II, type I/type III, and type II/type III. Although several genes are known to be involved during the sexual cycle (15), the determinant of mating type has proved elusive. Genetic analysis suggested that mating type is stable and is determined by a single locus with two or more alleles (10, 14, 16). We argued that it might be possible to identify this postulated locus by searching for genes that are present in any member of one mating type but absent or highly diverged in any member of another. For this purpose, we performed comparative genomic hybridizations using DNA microarrays composed of probes for around 8500 of the 10,500 predicted genes in the sequenced type I *D. discoideum* genome (17).

We analyzed 10 strains derived from independent wild isolates (table S1) using this micro-

¹Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. ²Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK. ³Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan.

*To whom correspondence should be addressed. E-mail: garethb@mrc-lmb.cam.ac.uk

MAUREEN MORRIN and S. RATNESHWAR*

Can pleasant ambient scents enhance consumer memory for branded products? If so, why? The authors examine the effects of ambient scent on recall and recognition of brands in two studies. In the first (i.e., encoding) phase of each study, subjects are asked to evaluate familiar and unfamiliar brands while viewing digital photographs of products on a computer screen; stimulus viewing times are measured covertly on the computer. Ambient scent is manipulated in the experiment room through a diffuser. In the second (i.e., retrieval) phase, conducted 24-hours later, brand recall and recognition accuracy are assessed. In both studies, ambient scent improves both recall and recognition of familiar and unfamiliar brands. This pattern emerges whether or not the scent is congruent with the product category (Study 1), and the enhancement in brand memory is due to the presence of ambient scent during encoding rather than retrieval (Study 2). Although ambient scent apparently did not alter subjects' self-assessed mood or arousal levels, it increased their attention in terms of longer stimulus viewing times. Mediation analyses suggest that the attention mechanism most likely explains why ambient scent improves brand memory.

Does It Make Sense to Use Scents to Enhance Brand Memory?

What is the association between scent and vivid memories of particular persons, products, or events from one's past? Anecdotal evidence has long hinted at a unique relationship between scent and human memory (Dichter 1998). Most famously, the French author Proust (1919), upon smelling a madeleine, was flooded with a series of childhood memories that became the basis for his classic novel *Remembrance of Things Past*. Scientific research generally supports the notion that humans can retain information about scents for long periods of time, perhaps all the way from childhood to late adulthood (Goldman and Seamon 1992; Laird 1935). Researchers are now beginning to understand why this is so (Holloway 1999). Still, at this point we know little about the influence of ambient scent on the processing of information and subsequent memory. This subject has generally received scant attention even in the basic disciplines such as cognitive psychology (Richardson and Zucco 1989).

*Maureen Morrin is Assistant Professor of Marketing, Rutgers University (e-mail: mmorrin@camden.rutgers.edu); this research was conducted while she was at the Katz Graduate School of Business, University of Pittsburgh. S. Ratneshwar is Professor of Marketing and Ackerman Scholar, School of Business, University of Connecticut (e-mail: ratti@business.uconn.edu). The authors thank Nada Nasr Bechwati for help in data collection. They are also grateful to Larry Barsalou and three anonymous *JMR* reviewers for their helpful comments and suggestions.

Despite the paucity of scientific research and theorizing, the practice of environmental fragrancing has grown into a major industry in recent years. Ambient scent is currently found in such diverse locations as retail stores (Helmsley 1997), supermarkets (Bainbridge 1998), restaurants (Petran 1998), office buildings (Marsh 1998), gambling casinos (Chase 1998), and even subway stations (*Global Cosmetic Industry* 1999). At least one company is currently testing in grocery stores the use of kiosks with touch-screen menus that are connected to scent-emitting devices (Schick 1999). Touching the key for a particular product results in the release of that product's scent into the atmosphere. Another firm is in the process of "scent-enabling the Internet" by selling computer hardware devices that attach to personal computers and emit scents according to specific software instructions (Wilder 1999). Such practices in the marketplace implicitly suggest that scents in the environment have the capacity to affect many aspects of consumer behavior.

Perhaps as a consequence, interest in the topic of ambient scent is finally growing among consumer researchers, and some key articles have appeared on this issue in recent years (Bone and Ellen 1999; Gulas and Bloch 1995; Knasko 1995; Mitchell, Kahn, and Knasko 1995; Spangenberg, Crowley, and Henderson 1996). For example, Spangenberg, Crowley, and Henderson (1996) examined the effects of ambient scent on in-store consumer evaluations of products in a simulated retail environment. They generally find that product evalua-

tions are more positive among subjects in scented environments. Furthermore, although subjects in the scented conditions did not actually spend significantly less or more time in the store than subjects in the unscented condition, they perceived themselves as having spent less time there.

Mitchell, Kahn, and Knasko (1995) took a different tack in that they examined the impact of ambient scent on consumer decision-making processes. Their results indicate that the presence of a congruent ambient scent (e.g., a floral scent while choosing floral arrangements), versus an incongruent scent (e.g., a chocolate scent in the same context), may increase the amount of time consumers spend making purchase decisions and the extent of variety-seeking behavior. Mitchell, Kahn, and Knasko also found that subjects in the congruent scent conditions produce more self-references and self-generated insertions (information about the product class not included in the stimuli) than subjects in the incongruent conditions.

These nascent findings suggest that ambient scent indeed affects a wide range of consumer behaviors. Notwithstanding, thus far relatively little attention has been devoted to the impact of scent on brand memory. This is quite unfortunate, because establishing a strong foothold in memory is an important strategic objective for most marketers of branded products (see, e.g., Aaker 1996; Keller 1993, 1998; Morrin 1999). The sheer accessibility of a brand name from memory has been shown to have a major impact on consumer choice (Hoyer and Brown 1990). Nedungadi (1990) demonstrates that increasing a brand's accessibility from memory by incidental priming substantially increases a brand's choice probability, even for familiar brands such as McDonald's. Furthermore, recent research suggests that for repeatedly purchased packaged goods, brand salience in memory may be the primary decision factor driving market share differences—outweighing even factors such as brand image or brand differentiation (Ehrenberg, Barnard, and Scriven 1997; Miller and Berry 1998).

Recognizing that there could be a link between ambient scent and brand memory, Morrin and Ratneshwar (2000) conducted a preliminary study in which they investigated the effects of scent on recall and recognition of real and fictitious brand names. They found that though a pleasant ambient scent increased the attention subjects paid to fictitious brand names, it had a weak and statistically nonsignificant effect on brand recall and no effect on brand recognition. It may be that Morrin and Ratneshwar's research suffered from a methodological shortcoming that undermined their findings. Specifically, in that study there was only a five-minute delay between the time of brand exposure and the time of brand retrieval. Therefore, it is possible that the effects of ambient scent on brand memory were not evident because subjects in general faced a relatively easy memory task. Conceivably, the beneficial effects of scent on stimulus attention would have manifested in brand memory if subjects had faced a more demanding task.

Therefore, the findings of Morrin and Ratneshwar (2000) are intriguing but equivocal in regard to the central question whether ambient scent can help create better memory for branded products. Their study also did not explore several related issues. For example, does it matter whether the ambient scent is congruent with the brand's product class? Does the scent need to be present at both encoding and retrieval of information to exert an effect on brand memory? Moreover,

even if ambient scents influence brand memory, what type of mechanism might account for this phenomenon? These are the critical questions that motivate the research reported here.

This article is organized as follows: We first present the conceptual background for the research and distinguish between various psychological processes that might account for any relationship between ambient scent and brand memory. We then report two experiments in which we examine this relationship empirically. We conduct these experiments in the context of an incidental learning task in which subjects are exposed to brand information through digital photographs of actual products on a computer screen. We manipulate ambient scent and include attention (stimulus viewing time), brand recall, brand recognition, and measures of mood and arousal as dependent variables. In accord with prior research, we assume that unpleasant scents have little practical relevance in marketing efforts and, therefore, examine only the impact of pleasant ambient scents.

CONCEPTUAL BACKGROUND

Mood and Arousal Mechanisms

Some researchers have suggested that ambient scent influences human behaviors such as work performance and productivity through mood mechanisms (see, e.g., Baron and Bronfen 1994; Baron and Thomley 1994). Still, although scents have been found in some studies to improve subjects' moods (e.g., Lawless 1991), the results have been decidedly mixed (e.g., Ellen and Bone 1998; Spangenberg, Crowley, and Henderson 1996). If ambient scents alter consumers' mood states, it is possible for ambient scent to have an impact on stimulus recall through *state-dependent memory mechanisms*. State-dependent memory refers to the finding that when a person's internal state (e.g., mood) at the time of encoding is reinstated at the point of retrieval, memory for the stimulus tends to be enhanced. Thus, studies have shown that when subjects learn and later attempt to retrieve information in the same physiological state (e.g., happy, sad, intoxicated), they are better able to remember such information than if they are in a different internal state (Bower 1981; Eich 1980). An extension of these findings is that similar results might be observed for brand memory if the presence of a pleasant ambient scent puts people in a happy mood. That is, if scent alters consumers' moods and state-dependent memory processes are at work, we should expect to find their memories improved, with a concomitant change in mood valence.

Another possibility is that the presence of an ambient scent increases consumers' physiological arousal levels and, consequently, brand memory. Indeed, some prior studies have shown that scents in the environment tend to increase physiological arousal levels (e.g., Donovan and Rossiter 1982; Lorig and Schwartz 1988). If the presence of ambient scent increases subjects' arousal levels, this should create higher levels of alertness, which in turn may result in better performance in cognitive tasks involving memory. Some indirect support for this line of reasoning is available in previous research. For example, Baron and Bronfen (1994) found that subjects completed more word construction and decoding tasks when air fresheners were added to the room's atmosphere, presumably because of increased physiological arousal. Furthermore, Keller and Block (1996)

have shown that increased physiological arousal results in greater cognitive elaboration. Taking both of these findings into consideration, we might expect to find that consumers' memories improve in the presence of ambient scent, with an accompanying change in physiological arousal levels.

Attention Mechanism

Another possibility is that ambient scents enhance performance in memory tasks on account of an attention rather than (or in addition to) mood and arousal mechanisms. Specifically, the presence of ambient scent may affect the encoding of information by facilitating "approach behaviors" and thus increasing the amount of attention consumers give to brand stimuli (Knasko 1995; Mehrabian and Russell 1974; Mitchell, Kahn, and Knasko 1995; Morrin and Ratneshwar 2000). Why might this occur? The sense of smell is believed by many researchers to act as a perceptual "gatekeeper" that helps an organism determine which environmental stimuli contribute to (or detract from) the overarching goal of survival (Goldstein 1996). Thus, people learn through experience that individual scents act as cues for either pleasant and approachable stimuli, such as food, or unpleasant stimuli that are better avoided, such as smoke and leaking gas (Goldstein 1996). As a result, pleasant ambient scents may increase consumers' approach behaviors, which should be evidenced by their inclination to linger in such scented environments.

Some evidence is available in prior research to support the notion that the presence of a pleasant ambient scent increases approach behavior among consumers. Knasko (1995) asked subjects to examine pictorial slides at their own pace. She found that the presence of a pleasant ambient scent resulted in longer viewing times. Furthermore, as discussed previously, Morrin and Ratneshwar (2000) found a similar effect with respect to fictitious brand names. Finally, even though Spangenberg, Crowley, and Henderson (1996) did not find that ambient scents increased the time spent by shoppers in a simulated store environment, they did find that it increased their approach behavior in terms of their intentions to visit the store in the future.

If a pleasant ambient scent induces approach behavior in that people take more time to view branded products in a scented environment, it is likely that such additional attention also causes the stimuli to be processed more thoroughly (see Fiske 1980; Huffman and Houston 1993; Pechmann and Stewart 1990). A robust set of findings in the memory literature suggests that the amount and depth of stimulus processing are highly correlated with the strength of memory traces and thus with the ease and accuracy of stimulus retrieval (e.g., Baddeley 1990; Craik and Tulving 1975). Therefore, additional attention at the point of stimulus processing should tend to create stronger, deeper memory traces, which later, at the time of attempted retrieval, should be evident in more accurate memory for stimulus brands. Note that prior researchers who examined the effects of ambient scent on attention either did not measure stimulus memory at all (e.g., Knasko 1995) or obtained mixed findings because of methodological limitations (e.g., Morrin and Ratneshwar 2000).

Summary: Alternative Process Mechanisms

Ambient scent in the atmosphere may result in superior memory for stimulus brands on account of three different

process mechanisms. First, a mood mechanism may be at work; if so, subjects in scented (versus control) conditions should display better memory and a more positive mood. Second, scent in the environment may enhance memory by elevating arousal and alertness. If this happens, subjects in scented (versus control) conditions should display better memory and higher arousal. Third, if scent enhances memory performance primarily through the route of increased attention, we should expect to find improved memory and longer stimulus viewing times among subjects in scented (versus control) conditions, without differences in measures such as mood and arousal. If more than one type of mechanism is responsible for the effect of ambient scent on memory, we would expect to observe a combination of these patterns.

The Role of Scent Congruity

Does the impact of ambient scent on brand memory depend on the extent of congruity between a particular scent and the product class of the stimulus brand? For example, are brands that are normally associated with floral scents better remembered if they are in an environment containing a floral scent rather than a food scent (see Mitchell, Kahn, and Knasko 1995)? This is one of the key issues we examine in this research. As mentioned previously, prior studies in this area have found some support for the notion that scent congruity is a relevant factor for certain consumer behaviors. However, those studies focused on issues and dependent variables quite different from the ones we employ in this research (see Ellen and Bone 1998, Mitchell, Kahn, and Knasko 1995). For example, Bone and Jantrania (1992) found that product-based odors that "fit" a product category, such as a cleaning solution with a lemon scent, improve product evaluations.

The basic premise underlying the expectation of a scent congruity effect on brand memory is one of semantic matching. Stronger links may be formed in memory when a brand is meaningfully related to an ambient scent. The associative network model of memory (Anderson 1983) implies that nodes in memory that share meaning (i.e., nodes that are highly semantically related) result in greater associative strength. Consequently, semantic matching between an ambient cue, such as scent, and brands encountered in that environment may lead to stronger associative bonds in memory. This idea is also consistent with the recent work of White and Treisman (1997) on serial position effects of odors, in which they find that scent-based memory may operate by people assigning specific verbal meanings to scent stimuli. With these possibilities in mind, we examine whether an ambient scent that is a better semantic match with brands encountered in the environment (i.e., a congruent scent) has a more beneficial impact on brand memory than one that is not (i.e., an incongruent scent).

SCENT PRETESTS

In selecting specific scents for Study 1, we sought to balance two important goals. First, we wanted to choose two scents that were quite disparate in regard to congruity with the focal product category of toiletry and household cleaning products that we planned to use in our main studies. Second, because prior research suggests that the most important dimension in human judgments of scents is hedonic preference (Ehrlichman and Halpern 1988), we needed

to select two scents that were more or less equally preferred so as to avoid a potential alternative explanation (i.e., confound) if any effects were to be observed on brand memory for the manipulation of scent congruity. Scent pretesting took place in two stages.

First Scent Pretest

We conducted the first scent pretest ($n = 27$) among undergraduates at a large northeastern university to determine which of four floral and woodsy scents was most preferred. We included three floral scents—rosemary, lavender, and geranium—that are generally rated favorably by humans (Moncrief 1970), along with a tree-based scent, eucalyptus. We believed that these scents could potentially serve as congruent scents for the category of toiletry and household cleaning products, and several of them have been used in prior research (e.g., Ludvigson and Rottman 1989; Morrin and Ratneshwar 2000; Spangenberg, Crowley, and Henderson 1996). The scents were identified by randomly assigned numbers, presented in small glass bottles containing a cotton ball with three to four drops of essential oil, and sniffed by pretest subjects in random order approximately six inches from the nose (we adapted this method from Spangenberg, Crowley, and Henderson's [1996] study). We then evaluated the scents on several nine-point semantic differential scales including measures of pleasantness (1 = "very unpleasant" to 9 = "very pleasant"), liking (1 = "do not like at all" to 9 = "like a lot"), and familiarity (1 = "not at all familiar" to 9 = "very familiar"). We averaged the pleasantness and liking scales to obtain a scent preference measure ($\alpha = .95$).

Figure 1 depicts descriptive statistics for the pretest data. A repeated-measures analysis of variance (ANOVA) con-

firmed that some scents were preferred more than others ($F(3, 78) = 4.58, p < .01, \eta^2 = 15\%$). Follow-up paired comparisons showed that geranium ($\bar{X} = 4.20$) was preferred more than the other three scents. It was rated more favorably than lavender ($\bar{X} = 2.62$; paired $t = 2.36, p < .01$; effect size Cohen's $d = .67$; see Cohen 1988) and rosemary ($\bar{X} = 3.14$; paired $t = 2.46, p < .05$; Cohen's $d = .43$) and directionally more favorably than eucalyptus ($\bar{X} = 3.29$; paired $t = 2.66, p < .10$; Cohen's $d = .34$). Therefore, we selected geranium as a reasonably well-liked and familiar floral scent that could potentially serve as the congruent scent in Study 1.

Second Scent Pretest

We conducted a second scent pretest ($n = 21$) to choose an ambient scent that would be as well liked as geranium but would be perceived as much less appropriate for toiletry and household cleaning products. Subjects provided their opinions about six different, unlabeled scents in a manner similar to that described in the first pretest. The six scents were geranium (i.e., the intended congruent scent), nutmeg, spearmint, cloves, tangerine, and cinnamon (all of the latter are food-based scents). We placed several drops of essential oil for each of these scents on cotton balls in glass jars. Subjects opened and sniffed the contents of the jars, in random order, and answered questions about each scent. We first asked the subjects whether they could identify the scent. They then rated their preference for each scent through ratings of liking and pleasantness on nine-point semantic differential scales, as in the first pretest ($\alpha = .98$). Finally, subjects rated the appropriateness (1 = "not at all appropriate" to 9 = "very appropriate") of each scent for each of four product categories—two toiletry and household cleaning product categories (underarm deodorant and window cleaner) and two food and beverage product categories (cookies and herbal teas).

Figure 2 depicts the data from this pretest. Geranium was perceived as fairly appropriate for toiletry and household cleaning products. Indeed, of the six scents in this pretest, only tangerine exhibited a slightly higher mean appropriateness rating, but the difference between geranium and tangerine was not statistically significant ($\bar{X} = 3.66$ versus $\bar{X} = 4.11$, not significant). Furthermore, the cloves scent was perceived as much less appropriate than geranium for toiletry and household cleaning products ($\bar{X} = 1.97$ versus $\bar{X} = 3.66, p < .01$; Cohen's $d = .71$) and as more appropriate for the food and beverage items ($\bar{X} = 3.80$ versus $\bar{X} = 2.47, p < .05$; Cohen's $d = .51$).

The preference ratings for geranium were in the midrange of the six scents and, most important, were consistent with our second goal: Subjects did not differ significantly in their preferences for the cloves scent versus the geranium scent ($\bar{X} = 4.71$ versus $\bar{X} = 3.98, p > .25$). Therefore, we chose the cloves scent as the incongruent ambient scent for Study 1. Note that we did not pick tangerine as the congruent scent instead of geranium because subjects preferred tangerine significantly more than cloves ($\bar{X} = 6.97$ versus $\bar{X} = 4.71, p < .005$).

STUDY 1

The main objectives of Study 1 were as follows: First, we wanted to investigate whether ambient scent would have any beneficial effects on brand memory. We included both unfamiliar and familiar brands in the study so as to assess the

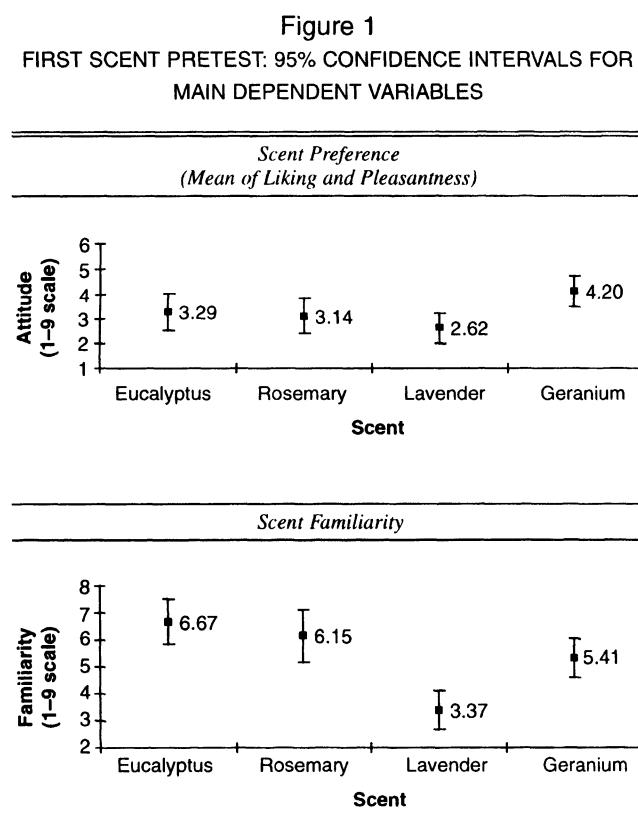
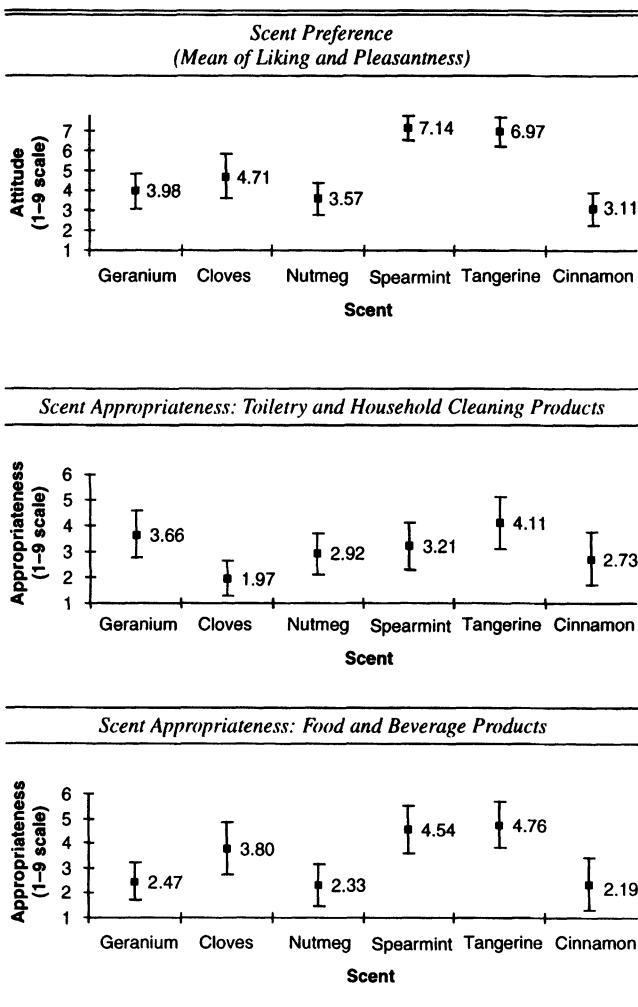


Figure 2

SECOND SCENT PRETEST: 95% CONFIDENCE INTERVALS
FOR MAIN DEPENDENT VARIABLES



generalizability of the findings. Second, if ambient scent has an effect on brand memory, we wanted to gain insights into the mechanisms that might be involved. Third, we wanted to explore whether scent congruity had any role in the relationship between ambient scent and brand memory.

Method

Design and subjects. The study consisted of a 3 (ambient scent) \times 2 (brand familiarity) mixed-model experiment. Ambient scent was a between-subjects factor, and 90 subjects were randomly assigned to one of three scent conditions: congruent scent, incongruent scent, or no scent. We exposed subjects assigned to a particular cell to the same ambient scent in the encoding and retrieval phases of the experiment. We manipulated brand familiarity as a within-subjects factor; all subjects were exposed to a set of unfamiliar and a set of familiar brand names. Students of an introductory business course at two large northeastern universities participated in the experiment in exchange for extra course credit. We included only native English speakers to exclude extraneous factors such as poor language fluency and varying levels of brand familiarity. Subjects signed up for two sessions of experiments, and the sessions were

spaced 24 hours apart. In reality, these two sessions were the encoding and retrieval phases of this study.

Brand stimuli. We chose our stimuli from the personal care, toiletry, and household cleaning product categories (e.g., deodorants, skin lotions, laundry detergents) because brands in these categories typically possess pleasant floral-based scents. These categories also have several brands with which we expected subjects to be familiar. A professional photographer took full-color digital photographs of all the stimulus products in their original packaging. These digital images served as our stimuli (for examples, see Figure 3).

We employed a total of 68 brands as stimuli in the experiment (for the full list, see the Appendix). We exposed subjects to half of these brands (i.e., the 34 target brands) during the encoding phase of the experiment. Half of the target brands (i.e., 17) were familiar brands for the subject population (e.g., Tide detergent, Ivory soap), and the other 17 brands were unfamiliar to the subjects (e.g., Margo soap, Happy Toes foot lotion). The unfamiliar brands were products sold in ethnic stores, small regional markets, or in parts of English-speaking Canada. In the recognition phase of the experiment, we exposed subjects to the target brands as well as to an equal number of foils (i.e., 17 familiar and 17 unfamiliar brands). Whenever possible, the foils were matched with the target brands in regard to product categories.

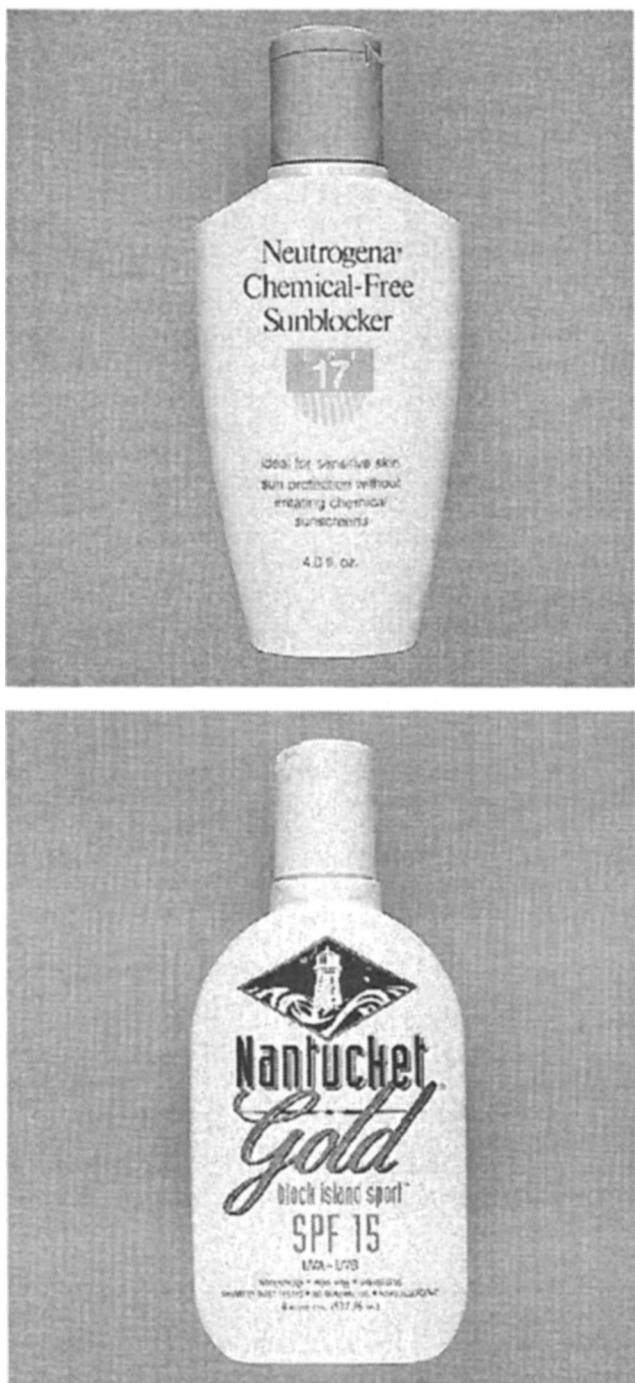
Procedure. In the scented conditions, the ambient scents were emitted into the room atmosphere by an electric diffuser into which several drops of the appropriate essential oil (i.e., geranium or cloves) had been placed. The diffuser emitted the scent into the atmosphere continuously during both the encoding and retrieval phases of the experiment. In the unscented condition, no scent was emitted. When running, the diffuser was placed behind a cardboard partition in a corner of the room so as to be quite inconspicuous. Although it was not possible to control the exact intensity of the scent emitted into the atmosphere, the same amount of essential oil was used in each session and the same amount of diffusion time was allowed to elapse before the commencement of the subject sessions for that day.

Subjects participated individually in the experiment. In the first (i.e., encoding) phase of the experiment, the subject entered either an unscented or a scented room (geranium or cloves) and completed a detailed screening questionnaire that assessed hunger, thirst, smoking habits, and allergies to scents (Engen 1982). Subjects with allergies or asthma would have been excused from participation; however, none of our subjects reported such problems. On account of the time involved in filling out the screening questionnaire, the subject experienced the scent (if present in the room) for several minutes before moving on to other tasks (see Cann and Ross 1989). Our goal was to allow enough time for the scent to potentially affect the subject, but not enough time for subjects to completely adapt to the scent (Dalton and Wysocki 1996). The subject then completed the standard Pleasure–Arousal–Dominance (PAD) scale so that we could assess his or her current mood and level of arousal (for more details, see Mehrabian and Russell [1974] and the “Dependent Measures” section).

After the subjects completed the PAD scale, we explained to them that the purpose of the experiment was to obtain their opinions about several brands, some of which would be familiar and some of which would be unfamiliar. We pro-

Figure 3

EXAMPLES OF FAMILIAR AND UNFAMILIAR BRAND STIMULI



vided subjects with an evaluative orientation for the task, and subjects were not aware that their memory for brand information would be tested later; thus, all of the learning was incidental rather than intentional (see Ratneshwar et al. 1997). Each subject was then seated in front of a computer and provided with detailed instructions on the screen. We informed the subject that pictures of several product brands

would appear on the monitor and that the task was to evaluate each brand in a self-paced manner. We told the subject to take as much or as little time as needed to evaluate each product. After several practice trials involving stimuli unrelated to the study, the subject was exposed to the 34 target brands, one by one, and in a random order. The digital image of the product package corresponding to each brand appeared and stayed on the screen until the subject provided an evaluation of the brand by hitting one of the numbered keys (1–9) at the top of the keyboard. A one-second delay was built in between the subject's key press for a particular trial and the onset of the next stimulus. The computer program captured both the brand rating and the time subjects took to view each stimulus until the point of key press. The subject was then excused for the day. Subjects left the laboratory with the belief that that particular experiment was over.

Twenty-four hours later (cf. Morrin and Ratneshwar 2000), the subject returned to the same room for the retrieval phase of the experiment. The scent condition in the retrieval phase for each subject was the same as in the encoding phase. In the retrieval phase, we administered a surprise free recall task to the subjects; each subject was given three minutes to list as many of the brands seen in the first (encoding) phase of the experiment as possible. Next, we gave the subject a computer-based brand recognition task. The task involved exposure to the digital photographs that corresponded to the same 34 target brands as in the encoding phase and, in addition, the photographs for all of the 34 foils. The photographs of the 68 brands appeared on the screen one by one and in a random order. The subject's task was to determine, as quickly but as accurately as possible, whether each brand was one of the brands seen in the first part of the experiment. The subject responded on each trial by hitting either the lower right "Y" key marked "yes" or the lower left "Z" key marked "no." We instructed subjects to keep their left and right index fingers poised over the respective keys before beginning the task. After each response, the subject was provided with feedback—when the subject responded correctly, a high-pitched beep sounded, and when the response was wrong, a low-pitched beep was heard. Providing such feedback is a fairly common method to encourage subjects in memory experiments to balance accuracy with speed in stimulus recognition tasks (Cameron and Frieske 1994).

After the brand recall and recognition tasks, the subject filled out a questionnaire. The first item was a manipulation check for brand familiarity. The remaining items assessed the subject's awareness of ambient scent in the room. We used a funneling technique for this purpose. The first question was open-ended: "While taking part in the experiment today, did you notice anything special about the room's atmosphere?" The subject was then asked to respond more specifically (i.e., "yes," "no," or "not sure") to the question: "Did you detect any odor or scent in the room?" Subjects then assessed the pleasantness of the room odor and indicated whether they believed the room odor, if any, affected them in any way during the experiment. We also included a check for hypothesis guessing. Subjects were then debriefed and thanked for their participation.

Dependent measures. We measured attention through stimulus viewing time—specifically, the number of mil-

liseconds a subject spent viewing a brand on the screen in the encoding phase before recording an evaluation on the computer. The attention data, which exhibited a typical long right tail in raw scores, were log transformed to create a more normal-shaped distribution before ANOVA (Howell 1992).

We assessed brand recall accuracy by the number of familiar and unfamiliar brands correctly recalled by a subject in the test phase (17 targets in each case). These recall data, which exhibited several extremely small numbers, were subjected to a square-root transformation before ANOVA to create a more normal-shaped distribution (Howell 1992). Note that such a transformation attenuates differences in raw scores between experimental conditions and thus implies a conservative data analysis strategy.

We assessed brand recognition accuracy for both familiar and unfamiliar brands by counting the number of hits (i.e., targets correctly recognized) minus the number of false alarms (i.e., foils wrongly recognized). Therefore, recognition accuracy scores were corrected for subject guessing as per conventional practice (see Baddeley 1990). A maximum score of 17 (i.e., 100%) for brand recognition implies that a subject correctly recognized all 17 target brands and there were no false alarms.

As previously noted, we measured subjects' mood and arousal states with the PAD scale (Mehrabian and Russell 1974). The PAD scale measures mood valence (i.e., "pleasure") and arousal with six semantic differential items for each construct, and it has been used often in prior ambient scent research (e.g., Ellen and Bone 1998; Knasko, Gilbert, and Sabini 1990; Spangenberg, Crowley, and Henderson 1996). Machleit and Eroglu (2000, p. 102) note that the PAD scale is the premier measure in the area of environmental psychology for assessing the impact of the environment on people. Furthermore, the PAD scale appears to be particularly well-suited for measuring mood and arousal as state variables in consumer behavior (e.g., Hui and Bateson 1991). After completing the PAD scale, subjects completed scales for brand evaluation (1 = "not at all favorable" to 9 = "extremely favorable"), brand familiarity (1 = "not at all familiar" to 9 = "extremely familiar"), and room odor pleasantness (1 = "not at all pleasant" to 9 = "extremely pleasant").

Results

We analyzed the main dependent measures with mixed-model ANOVAs. We conducted follow-up contrasts (two-tailed tests) when appropriate. The between-subjects independent variable, ambient scent, had three conditions (congruent scent, incongruent scent, or unscented control condition). Brand familiarity was a within-subjects variable and had two levels (unfamiliar and familiar brands). Table 1 shows means for all dependent measures for both the untransformed and the transformed data. However, to facilitate interpretation in the subsequent discussion and in Figure 4, we report means for the untransformed data.

Manipulation check for brand familiarity. As expected, familiar brands were rated as more familiar than unfamiliar brands ($\bar{X} = 7.57$ versus $\bar{X} = 1.98$; paired $t = 44.52$, $p < .0001$; Cohen's $d = 4.69$).

Brand attention. On the basis of our theoretical discussion, we anticipated that the presence of a pleasant ambient scent would increase approach behavior and therefore

increase brand attention. We also wanted to explore whether scent congruency would enhance this effect. The ANOVA on the attention data confirmed a main effect for ambient scent condition ($F(2, 87) = 3.29$, $p < .05$, $\eta^2 = 2.7\%$). Follow-up comparisons showed that subjects in the congruent scent condition gave the brand stimuli more attention than their counterparts in the unscented control condition ($\bar{X} = 5,421$ versus $\bar{X} = 4,107$ msec, $p < .0001$, $\eta^2 = 2.2\%$). Similarly, subjects in the incongruent scent condition also paid more attention to the stimuli than those in the unscented condition ($\bar{X} = 5,176$ versus $\bar{X} = 4,107$ msec, $p < .0001$, $\eta^2 = 1.8\%$). The difference between the congruent and incongruent scent conditions was not statistically significant ($p > .45$).

The attention data also yielded a main effect for brand familiarity ($F(1, 2967) = 280.4$, $p < .0001$, $\eta^2 = 5.3\%$). Subjects attended more to unfamiliar than familiar brands ($\bar{X} = 5,571$ versus $\bar{X} = 4,231$ msec), as might be expected on the basis of stimulus novelty (see Lynch and Srull 1982). Furthermore, the analysis revealed an interaction between scent condition and brand familiarity ($F(2, 2967) = 7.66$, $p < .0005$, $\eta^2 = 1\%$). Because the effect size of this interaction is extremely small when compared with the main effect for scent, we do not discuss it further.

Brand recall. We expected that the presence of a pleasant ambient scent would improve brand recall accuracy, and we were curious to know whether this effect would be especially evident when the scent was congruent with the product category. Confirming our expectations, the ANOVA showed a main effect for ambient scent condition ($F(2, 87) = 7.87$, $p < .001$, $\eta^2 = 3.1\%$). Follow-up comparisons revealed that subjects in the congruent scent (versus unscented) condition recalled more brands ($\bar{X} = 18\%$ versus $\bar{X} = 11\%$ of target brands correctly recalled, $p < .001$, $\eta^2 = 1.9\%$). A similar difference in brand recall was also found between the incongruent scent and unscented conditions ($\bar{X} = 18\%$ versus $\bar{X} = 11\%$, $p < .0001$, $\eta^2 = 2.7\%$). The contrast between the congruent and incongruent scent conditions was not statistically significant ($p > .45$). In addition, the main effect of brand familiarity was statistically significant ($F(1, 87) = 493.6$, $p < .0001$, $\eta^2 = 67.5\%$); not surprisingly, subjects recalled many more familiar than unfamiliar brands ($\bar{X} = 28\%$ versus $\bar{X} = 3\%$). The interaction between ambient scent condition and brand familiarity was not statistically significant ($F < 1$), and the anticipated pattern of results for the effects of scent held for both the unfamiliar and the familiar brands.

Brand recognition. Paralleling our expectations for brand recall, we anticipated that the presence of an ambient scent would improve brand recognition accuracy. The ANOVA indeed yielded a main effect for ambient scent condition ($F(2, 87) = 8.93$, $p < .0005$, $\eta^2 = 12.2\%$). Comparisons revealed that brand recognition accuracy was reliably higher in the congruent scent (versus unscented) condition ($\bar{X} = 80\%$ versus $\bar{X} = 67\%$, $p < .0001$, $\eta^2 = 10.5\%$). Brand recognition accuracy was also higher in the incongruent scent (versus unscented) condition ($\bar{X} = 78\%$ versus $\bar{X} = 67\%$, $p < .0001$, $\eta^2 = 7.7\%$). The difference between the congruent and incongruent conditions was not statistically significant ($p > .40$). Neither the main effect for brand familiarity ($F < 1$) nor the interaction between ambient scent condition and brand familiarity ($F(2, 87) = 1.20$, $p > .30$) was statistically significant.

Table 1
STUDY 1: THE EFFECTS OF AMBIENT SCENT ON BRAND ATTENTION AND MEMORY

<i>Dependent Variables</i>	<i>Ambient Scent Condition</i>		
	<i>No Scent (Control) (n = 30)</i>	<i>Congruent Scent (n = 30)</i>	<i>Incongruent Scent (n = 30)</i>
<i>Brand Attention (Stimulus Viewing Time)</i>			
All brands	4107 (8.18)	5421† (8.39)	5176† (8.37)
Unfamiliar brands	4498 (8.28)	6310† (8.57)	5906† (8.50)
Familiar brands	3716 (8.08)	4531† (8.22)	4446† (8.24)
<i>Brand Recall</i>			
All brands	11% (.99)	18%*** (1.35)	18%† (1.41)
Unfamiliar brands	0% (.13)	3%* (.45)	5%*** (.60)
Familiar brands	22% (1.85)	32%** (2.25)	31%** (2.23)
<i>Brand Recognition</i>			
All brands	67%	80%†	78%†
Unfamiliar brands	65%	81%†	78%†
Familiar brands	68%	77%**	77%**
<i>Brand Evaluation</i>			
All brands	4.76	5.18	4.90
Unfamiliar brands	3.77	4.18	4.01
Familiar brands	5.75	6.17	5.80
Mood	4.04	4.14	3.75
Arousal	4.80	4.84	4.59
Room odor pleasantness	5.50	5.50	5.55

Notes: The table shows cell means for all dependent variables, including those that were transformed before we conducted ANOVAs. Brand attention refers to stimulus viewing time in milliseconds, and the numbers in parentheses in the lower row refer to the same variable after a \log_n transformation. Brand recall is the percentage of stimulus brands correctly recalled, and the numbers in parentheses in the lower row refer to the square-root of the absolute number of brands correctly recalled. Brand recognition is the percentage of hits minus the percentage of false alarms. Brand evaluation, mood, arousal, and room odor pleasantness were assessed on one-to-nine scales, where higher numbers indicated more positive ratings and more arousal. See text for more details on all measures. Means within a row that are statistically significantly different from the no scent (control) condition are denoted as follows: * $p < .05$, ** $p < .01$, *** $p < .001$, and † $p < .0001$.

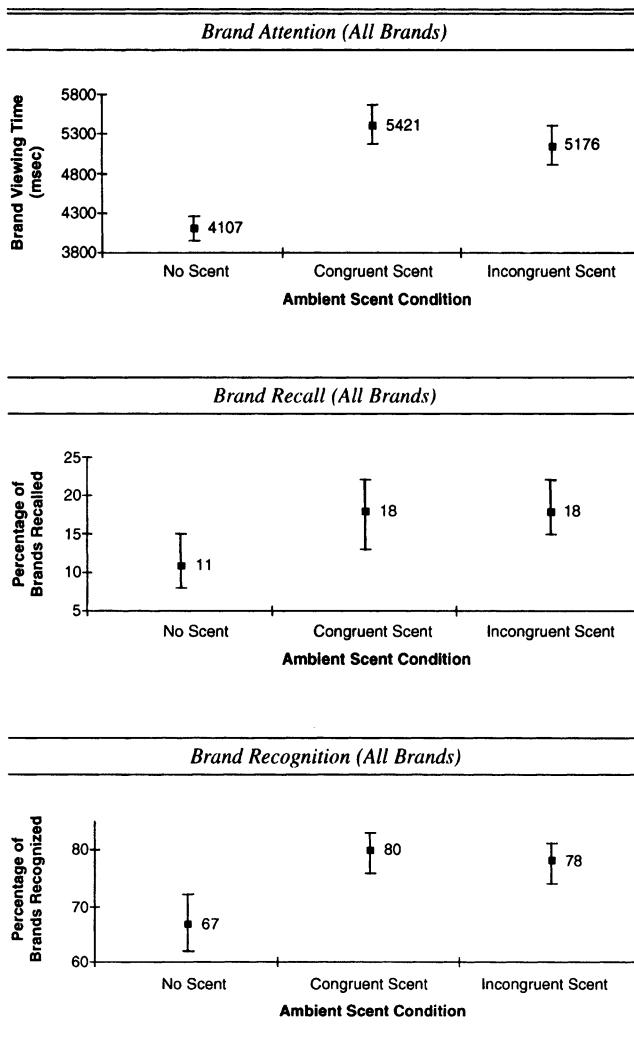
Mediation analyses. The foregoing results show that the presence of an ambient scent increased attention (a potential mediator) as well as brand recall and recognition accuracy (final outcome variables). If attention either partially or fully mediates the observed effects of scent on recall and recognition of brand stimuli, the inclusion of attention as an additional independent variable in the analyses of the two memory measures should result in a reduction of the main effect of ambient scent (Baron and Kenny 1986). Therefore, we performed additional analyses to verify whether this was the case. When the attention measure (stimulus viewing time) was added to the analysis for brand recall, the parameter for this independent variable did not achieve significance ($F < 1$). Thus, in the case of brand recall, we did not find much support for a mediating role for stimulus attention. Notwithstanding, when the attention variable was added to the model for brand recognition, it emerged as statistically significant ($F(1, 86) = 10.75, p < .005, \eta^2 = 9.6\%$). Furthermore, its inclusion in the model substantially reduced the effect size of ambient scent in regard to brand recognition ($\eta^2 = 12.2\%$ without the mediator versus 7.9% with the mediator). Taken as a whole, these results partially support the explanation that the influence of ambient scent on brand memory may involve attention as a key mediator.

Other dependent measures. We obtained brand evaluations in the encoding phase of the study. Although we had no *a priori* predictions, we conducted a mixed-model ANOVA on the brand ratings. Only the main effect of brand familiarity was statistically significant ($F(1, 2967) = 677.1, p < .0001, \eta^2 = 15.8\%$). As might be expected, subjects rated familiar brands more favorably than unfamiliar brands ($\bar{X} = 5.91$ versus $\bar{X} = 3.99$). Neither the main effect for scent condition ($F(2, 87) = 1.56, p > .20$) nor the ambient scent by brand familiarity interaction ($F < 1$) was statistically significant.

A one-way ANOVA of subjects' self-assessed mood ratings showed no effect for ambient scent condition ($F < 1$). A similar ANOVA revealed that the ambient scent condition had no impact on arousal ratings ($F < 1$). As can be seen from the means in Table 1, both mood and arousal ratings are uniformly in the midrange of the scales and do not vary much between scent conditions. These results suggest that the observed effects of ambient scent on attention and memory do not involve affective mechanisms such as mood or arousal. We return to this issue subsequently. Approximately half of the subjects in the scented conditions noticed "something special" about the room's atmosphere. (Note that at this point, they were still in the laboratory room with the

Figure 4

STUDY 1: 95% CONFIDENCE INTERVALS FOR MAIN DEPENDENT VARIABLES



scent diffuser running.) Most of the subjects who noticed something affirmed that it was the scent. But only one subject believed he was affected by the odor. Finally, an ANOVA on subjects' ratings of room odor pleasantness failed to show an effect for ambient scent condition ($F < 1$).

Discussion

The results of Study 1 provide important insights into whether and how ambient scent can influence brand memory. First, the presence of a pleasant ambient scent, whether congruent or incongruent with the product category, substantially increased subjects' attention to brand stimuli. Mean viewing times for the stimuli in the two scented conditions (versus the unscented control condition) increased by nearly 30%. Ambient scent also improved consumer memory for both familiar and unfamiliar brands, as assessed by recall and recognition measures. For example, accuracy in brand recall improved by 7% from 11% to 18%, (i.e., more than 60% over baseline performance), and brand recognition accuracy on average increased by about 12% in the scented conditions when compared with the unscented control condition.

Second, the incongruent ambient scent (i.e., cloves) and the congruent ambient scent (i.e., geranium) had about the same degree of impact on brand recall and recognition. We did not find any support for the idea that scent congruity has an important role in enhancing brand memory. Thus, it appears that the facilitating effects of ambient scent do not depend on a semantic matching process. Rather, the results for the attention measure and the partial support we obtained for the mediating role of attention suggest that pleasant ambient scents mainly induce people to remain longer in a given environment, thereby causing them to process brand stimuli in greater depth. We also note that our scent manipulations had no effect on subjects' self-ratings of mood and arousal, which implies that the impact of ambient scent on memory probably does not operate through these two routes, at least as measured by the PAD scale we used in this research.

Although Study 1 results represent noteworthy findings in regard to relationships among ambient scent, attention, and brand memory, the design of this experiment does not permit unambiguous answers to some important questions. Does ambient scent need to be present at *both* brand exposure and retrieval for improvement in brand memory? Or is it sufficient that ambient scent be present only during brand encoding? If increased attention at brand encoding alone accounts for all of the memory improvement, presumably the presence or absence of ambient scent at brand retrieval would be immaterial. However, as we point out in the next section, there are theoretical reasons to suggest the possibility that ambient scent may need to be present at both the retrieval and the encoding stages for beneficial effects on brand memory.

STUDY 2

The presence of ambient scent during both the encoding and retrieval phases may be critical on account of two different factors. We discuss each of these factors next to provide a conceptual backdrop for the design of Study 2.

The Role of Environmental Reinstatement

The well-known Encoding Specificity principle suggests that memory is improved if the contexts at encoding and retrieval are well matched (Tulving 1983). In terms of our research, such matching between the encoding and retrieval contexts may occur because of *environmental reinstatement* (Smith 1994). Environmental reinstatement involves the recreation of the learning environment at the time of test. In a classic study in this area, Godden and Baddeley (1975) manipulated the physical location—underwater or on land—where subjects learned information. They subsequently tested subjects' memory in either the same or the other environment. They find that recall is higher when the external environment is the same as that at the point of encoding.

Is it the case then that environmental reinstatement is critical for ambient scent to improve memory? A study by Schab (1990) is relevant to the present discussion. Schab (1990, Experiment 1) exposed subjects to 40 common adjectives with or without the presence of chocolate aroma in the room. Subjects were asked to generate antonyms for each of the adjectives on a self-paced basis. They were also told to think about and imagine the smell of chocolate at both learning and test. After a 24-hour delay, a free recall task was

administered. Post hoc comparisons showed that recall was indeed highest when subjects were exposed to scent at both learning and test. Schab (1990) explains his results in accord with the Encoding Specificity principle.

Schab's (1990) stimuli and experimental procedures were quite different from those we employed in Study 1. Notwithstanding, if ambient scent has an environmental reinstatement effect, it is plausible that the results obtained in Study 1 are not solely attributable to increased attention at the point of encoding. It could be argued, on the basis of the Encoding Specificity principle, that the presence of ambient scent at both encoding and retrieval played a vital part in improving brand memory in Study 1. In Study 2, by manipulating the presence versus absence of ambient scent at both the encoding and retrieval stages in a factorial design (as did Schab 1990), we shed light on this critical issue.

The Role of State Dependency

Although there was no evidence for a mood mechanism in Study 1's results, we believed it would be worth examining this issue again in Study 2. As discussed previously, researchers have found that information is better remembered if subjects learn and retrieve materials while in the same mood state. If state dependency is part of the explanation for the facilitating effects of ambient scent on memory, (1) recall and recognition levels should be highest when ambient scent is present at encoding and retrieval and (2) mood should be altered by ambient scent. Therefore, mood measures should enable us to distinguish state-dependency effects from the effects of encoding specificity and environmental reinstatement.

Method

This study consisted of a 2 (ambient scent present or absent at encoding) \times 2 (ambient scent present or absent at retrieval) \times 2 (brand familiarity) mixed-model experiment. We manipulated the ambient scent conditions between-subjects, and we randomly assigned subjects to one of the four experiment conditions. We manipulated brand familiarity as a within-subjects factor, as in Study 1. Sixty students of an introductory business course at a large northeastern university participated in this experiment in exchange for extra course credit. We included only native English speakers.

As in Study 1, subjects participated in two sessions (i.e., the encoding and retrieval phases) that were spaced 24 hours apart. The brand stimuli, procedure, and dependent measures employed were nearly identical to those of Study 1. Other than the manipulation of ambient scent in both the encoding and retrieval phases, the only other major difference in the Study 2 method was that scent congruency was not manipulated. Instead, only the congruent scent (i.e., geranium) of Study 1 was used in the scented conditions of Study 2. In addition, we probed scent awareness at the end of the experiment in a room other than the one in which the experiment took place, so that subjects' awareness levels would be based on their perceptions during the experiment.

Results

As in Study 1, we analyzed the main dependent measures using mixed-model ANOVAs. We conducted follow-up contrasts when appropriate. We transformed the attention (stimulus viewing time) and brand recall measures before the

ANOVAs, as in Study 1. Table 2 reports means for all dependent variables (transformed and untransformed), but as in Study 1, in the subsequent discussion and in Figure 5, we focus on the means for the untransformed data.

Manipulation check for brand familiarity. As expected, subjects rated familiar brands as more familiar than unfamiliar brands ($\bar{X} = 7.73$ versus $\bar{X} = 2.02$; paired $t = 39.8, p < .0001$; Cohen's $d = 5.15$).

Brand attention. We conducted a mixed-model ANOVA with attention as a function of the ambient scent condition at encoding, brand familiarity, and their interaction. (We did not include ambient scent at retrieval as an independent variable in this analysis because attention is relevant only in the encoding phase.) The main effect of ambient scent condition at encoding was statistically significant ($F(1, 58) = 8.62, p < .005, \eta^2 = 5.3\%$). As in Study 1, the presence (versus absence) of ambient scent at encoding enhanced attention; subjects in the scented condition, on average, took almost one second longer to view each brand ($\bar{X} = 4,079$ versus $\bar{X} = 3,104$ msec). The main effect of brand familiarity was also statistically significant; subjects paid more attention to unfamiliar (versus familiar) brands ($\bar{X} = 4,066$ versus $\bar{X} = 3,117$ msec; $F(1, 2148) = 240.3, p < .0001, \eta^2 = 5.9\%$). The ambient scent at encoding by brand familiarity interaction was not statistically significant ($F < 1$).

Brand recall. We conducted a mixed-model ANOVA on brand recall as a function of the ambient scent condition at encoding, scent condition at retrieval, brand familiarity, and all possible interactions. The main effect of scent condition at encoding was statistically significant ($F(1, 56) = 4.54, p < .05, \eta^2 = 1.4\%$). As in Study 1, and consistent with the foregoing findings on attention, subjects recalled more brands when an ambient scent was present (versus absent) at the point of encoding ($\bar{X} = 15\%$ versus $\bar{X} = 13\%$ of the target brands). As we anticipated, subjects also recalled many more of the familiar brands in comparison with the unfamiliar brands ($\bar{X} = 26\%$ versus $\bar{X} = 3\%$; $F(1, 56) = 233.61, p < .0001, \eta^2 = 63.4\%$). The presence (versus absence) of ambient scent at retrieval had no impact on brand recall ($F < 1$), and none of the interactions was statistically significant ($ps > .10$).

Brand recognition. We conducted a mixed-model ANOVA that was similar to the one on brand recall. Again, the results displayed a main effect for the ambient scent condition at encoding ($F(1, 56) = 5.57, p < .05, \eta^2 = 7\%$). Subjects' recognition accuracy was higher when brand encoding took place in an environment in which ambient scent was present rather than absent ($\bar{X} = 73\%$ versus $\bar{X} = 63\%$). As was the case for brand recall, we did not find a main effect for ambient scent at retrieval on brand recognition accuracy, and the interaction between scent at encoding and scent at retrieval was also not statistically significant ($F's < 1$). The analysis also revealed that brand familiarity did not influence recognition scores ($F < 1$), and none of the interactions in the model was statistically significant ($ps > .15$).

Mediation analyses. Similar to the findings of Study 1, the presence of ambient scent at encoding had a positive effect on attention (a potential mediator) and the two memory measures. Therefore, we performed mediation tests again to examine whether ambient scent at encoding improves recall and recognition through the route of attention. We added the attention mediator as another independent variable to the mixed model analyses of the memory

Table 2
STUDY 2: THE EFFECTS OF MANIPULATING AMBIENT SCENT AT BRAND ENCODING VERSUS RETRIEVAL

<i>Dependent Measures Retrieval</i>	<i>Scent Absent at Encoding</i>			<i>Scent Present at Encoding</i>		
	<i>Scent Absent at Retrieval (n = 15)</i>	<i>Scent Present at Retrieval (n = 15)</i>	<i>Margin Mean (n = 30)</i>	<i>Scent Absent at Retrieval (n = 15)</i>	<i>Scent Present at Retrieval (n = 15)</i>	<i>Margin Mean (n = 30)</i>
<i>Brand Attention (Stimulus Viewing Time)</i>						
All brands	3332 (7.96)	2914 (7.86)	3104 (7.91)	3760 (8.10)	4441 (8.23)	4079† (8.17)
Unfamiliar brands	3924 (8.12)	3228 (7.98)	3544 (8.05)	4181 (8.22)	5049 (8.37)	4588† (8.29)
Familiar brands	2739 (7.80)	2600 (7.74)	2663 (7.77)	3338 (7.99)	3832 (8.09)	3570† (8.04)
<i>Brand Recall</i>						
All brands	14% (1.15)	13% (.99)	13% (1.07)	15% (1.23)	15% (1.40)	15%* (1.31)
Unfamiliar brands	2% (.23)	1% (.13)	1% (.18)	3% (.38)	6% (.79)	4%*** (.59)
Familiar brands	27% (2.08)	24% (1.84)	26% (1.96)	28% (2.08)	25% (2.00)	26% (2.04)
<i>Brand Recognition</i>						
All brands	63%	63%	63%	73%	73%	73%†
Unfamiliar brands	66%	61%	64%	74%	74%	74%*
Familiar brands	59%	66%	63%	71%	73%	71%*
<i>Brand Evaluation</i>						
All brands	5.10	4.58	4.84	4.99	4.99	4.99
Unfamiliar brands	3.89	3.47	3.68	4.17	3.56	3.88
Familiar brands	6.32	5.69	5.99	5.81	6.41	6.09
Mood	4.76	3.82	4.29	4.19	4.09	4.14
Arousal	5.28	4.75	5.02	5.36	4.80	5.08
Room odor pleasantness	5.84	5.19	5.52	6.39	6.00	6.19

Notes: The table shows cell means for all dependent variables, including those that were transformed before we conducted ANOVAs. Brand attention refers to stimulus viewing time in milliseconds, and the numbers in parentheses in the lower row refer to the same variable after a \log_n transformation. Brand recall is the percentage of stimulus brands correctly recalled, and the numbers in parentheses in the lower row refer to the square-root of the absolute number of brands correctly recalled. Brand recognition is the percentage of hits minus the percentage of false alarms. Brand evaluation, mood, arousal, and room odor pleasantness were assessed on one-to-nine scales, where higher numbers indicated more positive ratings and more arousal. See text for more details on all measures. Margin means within a row that are statistically significantly different between the scent present at encoding and scent absent at encoding conditions are denoted as follows: * $p < .05$, ** $p < .01$, *** $p < .001$, and † $p < .0001$.

measures. In the case of brand recall, the attention variable was not statistically significant ($F < 1$). However, in the case of brand recognition, the attention parameter was statistically significant ($F(1, 55) = 8.14, p < .01, \eta^2 = 12.3\%$). Furthermore, the addition of attention to the model caused the effect size for ambient scent to drop substantially ($\eta^2 = 7\%$ without the mediator to 3.3% with the mediator). Therefore, as in Study 1, we found support for the mediating role of attention in the case of the memory measure (brand recognition) in which ambient scent had larger effects.

Other dependent measures. As in Study 1, we examined the brand evaluation data even though we had no a priori expectations. As in Study 1, the main effect of brand familiarity was statistically significant; subjects rated familiar brands more positively than unfamiliar brands ($\bar{X} = 6.05$ versus $\bar{X} = 3.78; F(1, 2148) = 658.12, p < .0001, \eta^2 = 20.9\%$). Neither the main effect of scent condition nor any of the interactions were statistically significant (F 's < 1).

We conducted ANOVAs to examine the effect of ambient scent condition at encoding on the mood and arousal measures. As in Study 1, we found that these two variables were not affected by ambient scent condition (both F 's < 1), and subjects were uniformly in the midrange of the mood and

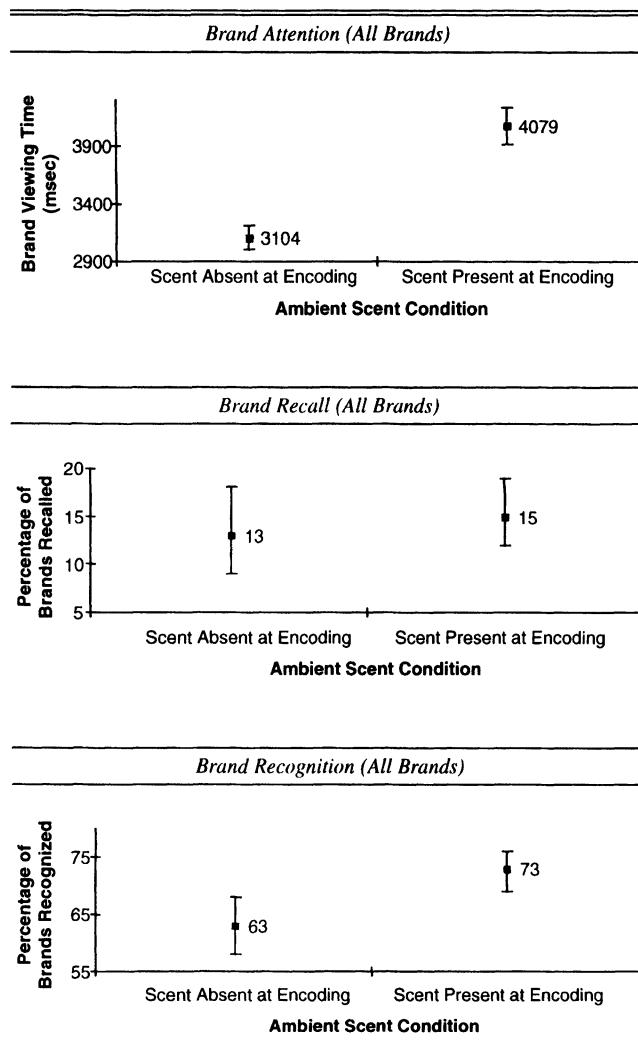
arousal scales. We conducted a 2 (scent condition at encoding) \times 2 (scent condition at retrieval) ANOVA on room odor ratings assessed at the end of the second phase of the study, but this analysis was not statistically significant ($p > .35$). Furthermore, relatively few of the subjects (17%) reported being aware of "anything special" about the room's atmosphere when queried at the end of the experiment.

Discussion

The findings from Study 2 provide corroboration for the results of Study 1 regarding the facilitating effects of ambient scent on brand recall and recognition. More important, the design of this experiment enabled us to evaluate more definitively whether attention at the point of brand exposure was the key factor driving the memory results or whether factors such as environmental reinstatement and state-dependency were implicated. The Study 2 data showed that brand memory was affected only by the presence (versus absence) of ambient scent at the brand encoding stage. Neither recall nor recognition was influenced by ambient scent at the brand retrieval stage. Furthermore, there was no evidence on either memory measure of an interaction between scent at encoding and scent at retrieval; this interaction

Figure 5

STUDY 2: 95% CONFIDENCE INTERVALS FOR MAIN DEPENDENT VARIABLES



should have been statistically significant if environmental reinstatement or state dependency contributed to the memory results. Finally, as we anticipated, ambient scent at encoding substantially increased attention (stimulus viewing time), and the addition of attention as an additional independent variable to the analysis of brand recognition diminished the independent effect of ambient scent on that memory measure. Therefore, as in Study 1, the results support the idea that ambient scent influences memory through the route of enhanced attention during brand encounters.

SUMMARY AND GENERAL DISCUSSION

Environmental fragrancing is now becoming a common practice in retailing, restaurants, and other service-oriented businesses. The impact of ambient scent on consumer behavior has also recently emerged as an important theoretical and empirical issue and has begun to attract the attention of scholarly researchers (e.g., Bone and Ellen 1999; Gulas and Bloch 1995; Mitchell, Kahn, and Knasko 1995; Morrin and Ratneshwar 2000; Spangenberg, Crowley, and Henderson 1996). Recent studies have investigated issues such as the effects of ambient scent on in-store product eval-

uations and consumer decision-making processes (e.g., Mitchell, Kahn, and Knasko 1995; Spangenberg, Crowley, and Henderson 1996).

Notwithstanding these developments, little is known about whether and why scents in the environment might influence consumer retrieval of brand information. This is an important gap in marketing knowledge in view of the well-established role of brand recall and recognition in consumer decision making (see, e.g., Hoyer and Brown 1990; Hutchinson, Raman, and Mantrala 1994; Keller 1993; Nedungadi 1990). We addressed this issue by examining the impact of pleasant ambient scents on brand memory in two experiments. We manipulated scent in the environment in the context of an incidental learning task wherein subjects viewed and evaluated branded products that were displayed sequentially on a computer screen. We measured stimuli viewing times covertly on the computer as a measure of attention, and we also obtained subjects' self-assessments of mood and arousal. Twenty-four hours later, we first assessed brand-free recall and then brand recognition in a discrimination task.

Across the two studies, we found a consistent pattern of results. When there was a pleasant ambient scent (versus no scent) during the brand exposure phase, subjects spent more time examining the stimuli. Concomitantly, in both studies, the presence of ambient scent improved subjects' recall and recognition of brand names. Additional analyses revealed that attention (i.e., stimulus viewing time) played a mediating role in enhancing brand memory.

Theoretical and Methodological Implications

Pleasant ambient scents appear to encourage approach behavior in that they induce consumers to spend more time in a particular environment.¹ Our finding of a positive relationship between a pleasant ambient scent and the time people spend in a given environment is consistent with several prior studies (Knasko 1995; Morrin and Ratneshwar 2000). However, Spangenberg, Crowley, and Henderson (1996) found virtually no effect for ambient scent on the time people spent in a simulated store environment, even though they hypothesized such an effect. Spangenberg, Crowley, and Henderson suggest that time spent in an environment is just one measure of approach behavior. Approach behavior can also be assessed in other ways, such as by a person's intention to spend more time in a given environment in the future. Spangenberg, Crowley, and Henderson found a positive effect for ambient scent on this latter measure of approach behavior. A possible explanation for the difference between our findings and their results may lie in the precision with which time is measured in the context of a research problem. Our research focuses on brand memory and measures stimulus viewing time through a computer in milliseconds; in contrast, Spangenberg, Crowley, and Henderson were primarily interested in product and store evaluations and measured time spent in a simulated store in seconds with a stopwatch.

In addition, an ambient scent that is congruent with a brand's product category did not result in differentially more attention or a greater improvement in brand memory than an incongruent scent (Study 1; see also Knasko 1995). This

¹We speculate that this may be a nonconscious process, in that most of the subjects in our studies were unaware of the presence of the ambient scent; those who were aware of the scent claimed to be unaffected by it.

suggests that pleasant ambient scents in general tend to improve memory performance and that scents that share semantic connotations with the products to be remembered do not necessarily result in deeper memory traces. The principal factor seems to be whether an ambient scent makes people stay longer in the environment such that they end up paying more attention to brand stimuli.

The results of Study 2 provide additional insights with respect to the mechanism responsible for improvement in brand memory. Specifically, the presence or absence of ambient scent at the point of brand retrieval had no discernable impact on memory performance; only the presence (versus absence) of ambient scent during brand encoding affected memory. Thus, our results do not support the theoretical possibility that ambient scent might improve memory by a process of environmental reinstatement (cf. Godden and Baddeley 1975; Schab 1990). Instead, the Study 2 data offer further evidence that ambient scent influences brand memory because it results in longer attention spans at the time of stimulus viewing, thereby creating deeper memory traces that are more easily retrieved, regardless of whether the scent is present again during brand retrieval. However, different results have been obtained in previous studies in which subjects expected a memory test and were made aware explicitly of the presence of an ambient scent during information encoding and retrieval (see, e.g., Schab 1990). We speculate that such procedures likely sensitize and alter subjects' information-processing strategies (see, e.g., Knasko, Gilbert, and Sabini 1990). However, with incidental learning tasks such as the one we created in our studies, the impact of ambient scent on memory probably operates through its effect on stimulus attention at the encoding phase, not through contextual cueing effects.

Ambient scent did not have an impact on subject's self-reports of mood and arousal in either study (for similar null results, see Spangenberg, Crowley, and Henderson 1996). Therefore, we did not find any evidence to suggest that these two variables are implicated in the results we observed for brand memory. To examine further the reliability of our null findings regarding mood and arousal, we conducted power analyses (Cohen 1988). When we collapsed the analyses across the two studies (which used identical procedures during the encoding phase of the experiment when mood and arousal were measured), assuming a two-tailed alpha of .05, we had an 85% probability of detecting a medium effect size for both mood and arousal. Therefore, we have a fair degree of confidence that the null results for these variables are not due to insufficient power.²

In contrast to our theorizing and findings, prior consumer researchers have often assumed an affective or emotional link between scent and behavior. For example, Ellen and Bone (1998, p. 29) suggest that "Odors stimulate the limbic system, the part of the brain responsible for emotional responses." We recommend that future researchers at least entertain the possibility that any observed effects of ambient scent on consumer behavior are the result of cognitive mechanisms such as attention and that they employ experimental methods that are capable of assessing this notion fur-

ther. In this regard, there are two noteworthy aspects to the present research methodology: (1) We employed digital photographs of actual branded products as stimuli, and (2) we measured consumer attention unobtrusively in an incidental learning task. Given the critical role of attention in our brand memory findings, it is nontrivial that we presented ecologically valid stimuli to our subjects rather than mere brand names (cf. Morrin and Ratneshwar 2000). The generalizability of our findings to real-world settings would have been questionable if we had limited our research to sterile verbal stimuli (Klink and Smith 2001). Furthermore, although previous researchers have used reaction-time measures to assess the strength of brand associations in memory (e.g., Herr, Farquhar, and Fazio 1996), we employed computer measures of brand viewing time during an incidental exposure task. Our methodology may be useful for other consumer researchers interested in using unobtrusive but precise measures of attention (see also Pechmann and Stewart 1990).

A few other theoretical aspects of our memory results are worthy of comment. In both studies, subjects were much more accurate in brand recognition than in brand recall. This finding replicates previous research in memory that has shown that the difference in accuracy between recognition and free recall is usually quite high in incidental learning tasks (see Eagle and Leiter 1964). Furthermore, although not the focus of our research, we note that subjects (1) attended much longer to unfamiliar brands (presumably on account of stimulus novelty), (2) nevertheless recalled familiar brands far better than unfamiliar brands, but (3) recognized familiar and unfamiliar brands about equally well. That brand familiarity facilitates brand recall but not brand recognition has a parallel in the well-known word frequency \times memory task interaction in the memory literature (e.g., Kinsbourne and George 1974). Previous researchers in the field of cognitive psychology have found this pattern in tasks in which subjects are given a fixed amount of time to process both familiar and unfamiliar stimuli and in intentional learning situations in which subjects anticipate a memory test. Our results extend this research by showing that this interaction also occurs in an incidental learning task and that it occurs notwithstanding that subjects pay more attention to unfamiliar stimuli.

Limitations

The research presented here has some of the same limitations as most theory-testing experimental investigations in that we conducted the studies in a controlled laboratory setting and with a convenience sample of student subjects. The scents and brand stimuli we employed were also somewhat limited in nature. Scents other than the ones we used (i.e., geranium and cloves) may operate differently in their ability to enhance attention and memory or in their impact on mood and arousal. Further research could examine the generalizability of the present findings to other kinds of scents (e.g., food smells). We also confined our brand stimuli to the types of products that typically contain scent as an ingredient. Therefore, we do not know whether our findings would apply to product categories that are usually not scented (e.g., clothing, consumer durables). However, given that congruity between ambient scent and the product category failed to emerge as an important factor in Study 1, we speculate that our findings may well carry over to nonscented products.

²If we remove the 30 subjects from Study 1 who were exposed to the cloves scent, thus including only control subjects and those who were exposed to the geranium scent at encoding, the power to detect a medium effect size is reduced only marginally from .85 to .78.

Regarding the processes by which ambient scent might influence brand memory, as noted previously, we did not find any evidence to suggest that mechanisms such as mood or arousal are at work. However, it could be argued that this null finding might simply be due to the self-report measuring instrument we employed. Perhaps the PAD scale is not sensitive enough to detect the small alterations in mood and/or arousal that are brought about by ambient scent (for discussions of alternative measures of consumer mood and emotions, see Machleit and Eroglu 2000; Richins 1997). However, as we discussed in the methodology section, there is considerable precedent for using the PAD scale for measuring mood and arousal as state variables. It is left for further research to ascertain whether the present findings hold even when other procedures are used for measuring these variables (i.e., other mood and arousal measures or physiological indicators instead of paper-and-pencil self-reports).

Also, our research did not examine the role of individual differences (Knasko 1992). Some prior research suggests that people who are extremely shy exhibit lower olfactory thresholds; that is, they are better able to detect the presence of ambient scent (Herbener, Kagan, and Cohen 1989). Therefore, it is possible that individual differences not captured in the present research moderate the effects of ambient scent on attention and brand memory. Future studies may want to explore potential moderating variables such as olfactory thresholds, individual preferences for scent, optimal stimulation level, gender, and a variety of personality variables. Further research may also reveal that groups of consumers can be segmented according to scent preferences or styles.

Another issue not explored in the present research is whether informing consumers of the presence of scent alters its impact on attention and memory processes. Research in the person perception area suggests that attribute priming effects are seriously diminished when subjects become conscious of obtrusive priming manipulations (see Higgins, Bargh, and Lombardi 1985). If further research shows that a similar result occurs for ambient scent interventions, it could raise public policy issues regarding the need to notify consumers in retail and other environments in which ambient scent is used.

Managerial Implications

As noted previously, marketers and retailers employ environmental fragrancing in a wide variety of circumstances, from in-store use to in conjunction with the Internet. In this regard, our research supports previous findings in the literature that pleasant ambient scents can induce consumer approach behavior (Knasko 1995; Spangenberg, Crowley, and Henderson 1996). But our research substantially extends prior research by clearly demonstrating that consumers are likely to devote more attention to stimuli encountered in pleasantly scented environments. Therefore, we opine that ambient scents can be used in a variety of venues such as retail stores, automobile showrooms, trade shows, and product demonstrations to increase not only the amount of time spent in a given environment but also the amount of mental attention devoted to relevant stimuli (e.g., products, brand names) encountered in that environment.

Even more important, perhaps, is our finding that the causal chain between ambient scent and enhanced brand attention also extends to improved brand memory. More-

over, congruency of the ambient scent with the product category does not seem to be a relevant factor in memory improvement. If a marketer's goal is to induce consumers to pay more attention to and remember better the products and brands they see in a particular setting, the use of virtually any pleasant ambient scent should achieve this objective. Thus, a firm selling beauty products and toiletries does not necessarily need to choose floral scents; a bakery need not limit itself to only vanilla, cinnamon, or other scents and spices typically used in baked goods.

Our findings should be highly relevant to brand marketers who set up and pay for product displays and promotions in retail stores and other venues. For example, adding scent-emitting devices to in-store displays and "shelf-talkers" could substantially enhance subsequent brand awareness and brand recall. If we extrapolate these results even further, it is possible that if devices such as scent strips are added to print advertisements or to free-standing coupon inserts in newspapers, consumers may bestow more attention to the advertised brand and thereby recall the advertised brand better at the point of purchase. Nevertheless, given that the effect sizes we obtained were somewhat small in some cases (especially in regard to brand recall), the cost-benefit aspects of such tactics will need to be thoroughly evaluated before they are implemented. It also remains to be seen whether ambient scent can be a cost-efficient device for enhancing brand memory when compared with traditional marketing tactics such as repetition in media advertising, product placement in movies, sponsorship of televised sports events, and so forth.

Conclusion

Taken as a whole, this research suggests a yes to the question posed in the title of this article—it does make sense to use scents to enhance brand memory. More generally, our work adds to the small but growing body of knowledge on whether, how, and when scents in the environment are likely to affect different facets of consumer behavior such as memory, judgments, decisions, and consumption behaviors. Our research also demonstrates that for studying the effects of ambient scent in marketing and consumer research settings, theory and methodological tools from cognitive psychology can be successfully adapted and applied to brand and product stimuli that have considerable ecological validity. Further research on ambient scent that builds on these beginnings should advance theorizing regarding the role of contextual factors in consumer behavior and also offer an even wider range of managerial suggestions.

REFERENCES

- Aaker, David A. (1996), *Building Strong Brands*. New York: The Free Press.
- Anderson, John R. (1983), "A Spreading Activation Theory of Memory," *Journal of Verbal Learning and Verbal Behavior*, 22 (3), 261-95.
- Baddeley, Alan (1990), *Human Memory: Theory and Practice*. Boston: Allyn and Bacon.
- Bainbridge, Jane (1998), "Scenting Opportunities," *Marketing*, (February 19), 36-37.
- Baron, Reuban M. and David A. Kenny (1986), "The Moderator-Mediator Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173-82.
- Baron, Robert and Marna I. Bronfen (1994), "A Whiff of Reality: Empirical Evidence Concerning the Effects of Pleasant Fra-

Appendix
BRAND STIMULI USED IN STUDIES 1 AND 2

	<i>Target Brands</i>	<i>Foil Brands</i>
Familiar brands	Tide detergent Edge shave cream Snuggle fabric softener Camay soap Crest toothpaste Close-Up toothpaste Gillette aftershave Sure deodorant Wizard air freshener Pine Sol cleanser Neutrogena sunblock Joy dish detergent Dawn dish detergent Ivory soap Suave shampoo/conditioner VO5 shampoo Finesse conditioner	Purex detergent Barbasol shave cream Bounce fabric softener Caress soap Colgate toothpaste Johnson & Johnson dental floss Rise shave cream Degree deodorant Glade air freshener Lysol cleanser Intensive Care sunblock Sunlight dish detergent Palmolive dish detergent Jergens soap Swiss Formula deodorant Pantene shampoo Vidal Sassoon conditioning rinse
Unfamiliar brands	Pears skin lotion Dessert Essence skin lotion Magical Mane conditioner Liril talc Argo laundry starch Germ-X soap Green Mark dishwashing detergent Wright's silver polish Clearly Natural soap Orchard Blend skin lotion Herbissimo shower gel California Scents air freshener Margo soap Nantucket Gold sunblock Smells Be Gone deodorizing spray Citrus Magic hand soap Thicker Fuller Hair styling gel	Herbal Mint skin lotion Aloe Vera skin lotion European Mystique shampoo Happy Toes foot cream Fels-Naptha laundry soap Germ Buster soap Sun and Earth cleaner Gonzo stain remover Emlin soap Healthy Salon skin lotion Healing Garden shower gel Country Treasures air freshener Forest Pure soap Native Tan sunblock Mango Madness body spray Zud cleanser Fruit of the Earth skin lotion

- grances on Work-Related Behavior," *Journal of Applied Social Psychology*, 24 (13), 1179–1203.
- and Jill Thomley (1994), "A Whiff of Reality: Positive Affect as a Potential Mediator of the Effects of Pleasant Fragrances on Task Performance and Helping," *Environment and Behavior*, 26 (6), 766–84.
- Bone, Paula Fitzgerald and Pam Scholder Ellen (1999), "Scents in the Marketplace: Explaining a Fraction of Olfaction," *Journal of Retailing*, 75 (2), 243–62.
- and Swati Jantrania. (1992), "Olfaction as a Cue for Product Quality," *Marketing Letters*, 3 (3), 289–96.
- Bower, Gordon H. (1981), "Mood and Memory," *American Psychologist*, 36 (2), 129–48.
- Cameron, Glen T. and David A. Frieske (1994), "The Time Needed to Answer: Measurement of Memory Response Latency," in *Measuring Psychological Responses to Media*, Annie Lang, ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 149–64.
- Cann, Arnie and Debra A. Ross (1989), "Olfactory Stimuli as Context Cues in Human Memory," *American Journal of Psychology*, 102 (1), 91–102.
- Chase, Victor D. (1998), "Making Stereophonic Scents," *Appliance Manufacturer*, 46 (1), 12.
- Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Craik, Fergus I.M. and Endel Tulving (1975), "Depth of Processing and Retention of Words in Episodic Memory," *Journal of Experimental Psychology: General*, 104 (3), 268–94.
- Dalton, Pamela and Charles J. Wysocki (1996), "The Nature and Duration of Adaptation Following Long-Term Odor Exposure," *Perception & Psychophysics*, 58 (5), 781–92.
- Dichter, Peter (1998), "My Nose Remembers," *Global Cosmetic Industry*, 163 (3), 54–56.
- Donovan, Robert and John Rossiter (1982), "Store Atmosphere: An Environmental Psychology Approach," *Journal of Retailing*, 58 (Spring), 34–57.
- Eagle, Morris and Eli Leiter (1964), "Recall and Recognition in Intentional and Incidental Learning," *Journal of Experimental Psychology*, 68 (1), 58–63.
- Ehrenberg, Andrew, Neil Barnard, and John Scriven (1997), "Differentiation or Salience," *Journal of Advertising Research*, 37 (6), 7–14.
- Ehrlichman, Howard and Jack N. Halpern (1988), "Affect and Memory: Effects of Pleasant and Unpleasant Odors on Retrieval of Happy and Unhappy Memories," *Journal of Personality and Social Psychology*, 55 (5), 769–79.
- Eich, James E. (1980), "The Cue-Dependent Nature of State-Dependent Retrieval," *Memory & Cognition*, 8 (2), 157–73.
- Ellen, Pam Scholder and Paula Fitzgerald Bone (1998), "Does It Matter If It Smells? Olfactory Stimuli as Advertising Executive Cues," *Journal of Advertising*, 27 (4), 29–39.
- Fiske, Susan T. (1980), "Attention and Weight in Person Perception: The Impact of Negative and Extreme Behavior," *Journal of Personality and Social Psychology*, 38 (6), 889–906.
- Global Cosmetic Industry (1999), "The Scent of the Subway," 164 (3), 10.
- Godden, Duncan and Alan D. Baddeley (1975), "Context-Dependent Memory in Two Natural Environments: On Land and Under Water," *British Journal of Psychology*, 66 (3), 325–31.
- Goldman, William P. and John G. Seamon (1992), "Very Long-Term Memory for Odors: Retention of Odor-Name Associa-

- tions," *American Journal of Psychology*, 105 (4), 549–63.
- Goldstein, E. Bruce (1996), *Sensation and Perception*, 4th ed. Pacific Grove, CA: Brooks/Cole Publishing Co.
- Gulas, Charles S. and Peter H. Bloch (1995), "Right Under Our Noses: Ambient Scent and Consumer Responses," *Journal of Business and Psychology*, 10 (1), 87–98.
- Helmsley, Steve (1997), "Scents and Sensibility," *Marketing Week*, 20 (34), 45–50.
- Herbener, Ellen S., Jerome Kagan, and Matthew Cohen (1989), "Shyness and Olfactory Threshold," *Personality and Individual Differences*, 10 (11), 1159–63.
- Herr, Paul M., Peter H. Farquhar, and Russell H. Fazio (1996), "Impact of Dominance and Relatedness on Brand Extensions," *Journal of Consumer Psychology*, 5 (2), 135–59.
- Higgins, E. Tory, John A. Bargh, and Wendy Lombardi (1985), "Nature of Priming Effects on Categorization," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11 (January), 59–69.
- Holloway, Marguerite (1999), "The Ascent of Scent," *Scientific American*, 281 (5), 42–44.
- Howell, David C. (1992), *Statistical Methods for Psychology*, 3d ed. Belmont, CA: Duxbury Press.
- Hoyer, Wayne D. and Steven P. Brown (1990), "Effects of Brand Awareness on Choice for a Common, Repeat-Purchase Product," *Journal of Consumer Research*, 17 (September), 141–48.
- Huffman, Cynthia and Michael J. Houston (1993), "Goal-Oriented Experiences and the Development of Knowledge," *Journal of Consumer Research*, 20 (2), 190–217.
- Hui, Michael K. and John E.G. Bateson (1991), "Perceived Control and the Effects of Crowding and Consumer Choice on the Service Experience," *Journal of Consumer Research*, 18 (September), 174–84.
- Hutchinson, J. Wesley, Kalyan Raman, and Murali K. Manrala (1994), "Finding Choice Alternatives in Memory: Probability Models of Brand Name Recall," *Journal of Marketing Research*, 31 (November), 441–61.
- Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.
- (1998), *Strategic Brand Management*. Upper Saddle River, NJ: Prentice Hall.
- Keller, Punam Anand and Lauren Goldberg Block (1996), "Increasing the Persuasiveness of Fear Appeals: The Effect of Arousal and Elaboration," *Journal of Consumer Research*, 22 (4), 448–60.
- Kinsbourne, Marcel and James George (1974), "The Mechanism of the Word-Frequency Effect on Recognition Memory," *Journal of Verbal Learning and Verbal Behavior*, 13 (1), 63–69.
- Klink, Richard R. and Daniel C. Smith (2001), "Threats to External Validity of Brand Extension Research," *Journal of Marketing Research*, 38 (August), 326–35.
- Knasko, Susan C. (1992), "Ambient Odor's Effect on Creativity, Mood, and Perceived Health," *Chemical Senses*, 17 (1), 27–35.
- (1995), "Pleasant Odors and Congruency: Effects on Approach Behavior," *Chemical Senses*, 20 (5), 479–87.
- , Avery N. Gilbert, and John Sabini (1990), "Emotional State, Physical Well-Being, and Performance in the Presence of Feigned Ambient Odor," *Journal of Applied Social Psychology*, 20 (16), 1345–57.
- Laird, Donald A. (1935), "What Can You Do with Your Nose?" *Scientific Monthly*, 41, 126–30.
- Lawless, Harry T. (1991), "A Sequential Contrast Effect in Odor Perception," *Bulletin of the Psychonomic Society*, 29 (4), 317–19.
- Lorig, Tyler and Gary E. Schwartz (1988), "Brain and Odor: I. Alteration of Human EEG by Odor Administration," *Psychobiology*, 16 (3), 281–84.
- Ludvigson, H. Wayne and Theresa R. Rottman (1989), "Effects of Ambient Odors of Lavender and Cloves on Cognition, Memory, Affect, and Mood," *Chemical Senses*, 14 (4), 525–36.
- Lynch, John G., Jr., and Thomas K. Srull (1982), "Memory and Attentional Factors in Consumer Choice: Concepts and Research Methods," *Journal of Consumer Research*, 9 (June), 18–37.
- Machleit, Karen and Sevgin A. Eroglu (2000), "Describing and Measuring Emotional Response to Shopping Experience," *Journal of Business Research*, 49 (2), 101–11.
- Marsh, Brenda (1998), "No Scents Is Good Sense," *OH&S Canada*, 14 (1), 26–27.
- Mehrabian, Albert and James A. Russell (1974), *An Approach to Environmental Psychology*. Cambridge, MA: The MIT Press.
- Miller, Stephen and Lisette Berry (1998), "Brand Salience Versus Brand Image: Two Theories of Advertising Effectiveness," *Journal of Advertising Research*, 38 (5), 77–82.
- Mitchell, Deborah J., Barbara E. Kahn, and Susan C. Knasko (1995), "There's Something in the Air: Effects of Congruent or Incongruent Ambient Odor on Consumer Decision Making," *Journal of Consumer Research*, 22 (2), 229–38.
- Morrin, Maureen (1999), "The Impact of Brand Extensions on Parent Brand Memory Structures and Retrieval Processes," *Journal of Marketing Research*, 36 (November), 517–25.
- and S. Ratneshwar (2000), "The Impact of Ambient Scent on Evaluation, Attention, and Memory for Familiar and Unfamiliar Brands," *Journal of Business Research*, 49 (2), 157–65.
- Nedungadi, Prakash (1990), "Recall and Consumer Consideration Sets: Influencing Choice Without Altering Brand Evaluations," *Journal of Consumer Research*, 17 (December), 263–76.
- Pechmann, Cornelia and David W. Stewart (1990), "The Effects of Comparative Advertising on Attention, Memory, and Purchase Intentions," *Journal of Consumer Research*, 17 (2), 180–91.
- Petran, Meredith (1998), "Bakes Great, Less Filling," *Restaurants & Institutions*, 108 (8), 100.
- Proust, Marcel (1919), *A La Recherche Du Temps Perdu*. Paris: Gallimard.
- Ratneshwar, S., Luk Warlop, David G. Mick, and Gail Seeger (1997), "Benefit Salience and Consumers' Selective Attention to Product Features," *International Journal of Research in Marketing*, 14 (July), 245–60.
- Richardson, John T. and Gesualdo M. Zucco (1989), "Cognition and Olfaction: A Review," *Psychological Bulletin*, 105 (3), 352–60.
- Richins, Marsha (1997), "Measuring Emotions in the Consumption Experience," *Journal of Consumer Research*, 24 (2), 127–46.
- Schab, Frank R. (1990), "Odors and the Remembrance of Things Past," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16 (4), 648–55.
- Schick, Shane (1999), "NCR's Sense of Smell Sells," *Computer Dealer News*, 15 (43), 42.
- Smith, Steven M. (1994), "Theoretical Principles of Context-Dependent Memory," in *Theoretical Aspects of Memory*, 2d ed., Peter Morris and Michael Gruneberg, eds. London: Routledge, 168–95.
- Spangenberg, Eric R., Ayn E. Crowley, and Pamela W. Henderson (1996), "Improving the Store Environment: Do Olfactory Cues Affect Evaluations and Behaviors?" *Journal of Marketing*, 60 (April), 67–80.
- Tulving, Endel (1983), *Elements of Episodic Memory*. Oxford, UK: Oxford University Press.
- White, Theresa L. and Michel Treisman (1997), "A Comparison of the Encoding of Content and Order in Olfactory Memory for Visually Presented Verbal Materials," *British Journal of Psychology*, 88 (3), 459–69.
- Wilder, Clinton (1999), "Technology That Makes Scents," *Informationweek*, (November 8), 17.

Effects of Perceptual Uncertainty on Arousal and Preference Across Different Visual Domains

AQ: 1-2

AQ: au

Thomas Z. Ramsøy
and Morten Friis-Olivarius
Copenhagen Business School and Copenhagen University Hospital Hvidovre

Catrine Jacobsen and Simon B. Jensen
Copenhagen Business School

Martin Skov
Copenhagen Business School and Copenhagen University Hospital Hvidovre

To what extent can simple contextual events affect preference? In this study, three tests were applied to assert whether contextual unpredictability has a negative effect on preference for novel visual items. By asking subjects to rate their first impressions of novel brand logos while playing simple sounds, Study 1 shows that brand logos coupled to unpredictable sounds were rated less favorably than logos presented with a predictable sound. In Study 2, this effect is found to be equally strong for abstract art paintings. Finally, Study 3 demonstrates that the negative effect of unpredictable sounds on preference is associated with a stronger arousal response, as indexed by pupil dilation responses. These results suggest that unpredictable sounds engage an emotional response that affects the first impression of a concurrently presented visual object. We discuss these findings in light of the basic psychology and neuropsychology of preference formation.

Keywords: first impression, neuroscience, pupil dilation, arousal, emotion, preference formation

Preferences are well known to be affected by contextual factors. Indeed, the very foundation of prominent theories of decision making, such as Behavioral Decision Theory, has been the realization that rationality is bounded, that is, that preferences and decisions are influenced by contextual factors. In particular, factors such as risk and ambiguity have been shown across different situations to affect reported and revealed preferences (Kahneman & Tversky,

1979; Christopoulos et al., 2009). For example, gambles with known risks are preferred to gambles with unknown (ambiguous) risks, a factor that has been related to increased engagement of brain regions associated with aversion responses (Levy et al., 2010). Similarly, ambiguity aversion has been shown to systematically lead to a favoring of established brands to novel brands (Muthukrishnan, Wathieu & Xu, 2009). Although studies of preference and decision making treat uncertainty and ambiguity as higher mental processes—as found in financial decisions (Critchley, Mathias & Dolan, 2001), gambling (Bechara, Damasio & Damasio, 2000; Fukui et al., 2005), and multiattribute product choices (Muthukrishnan, Wathieu & Xu, 2009)—studies have also suggested that similar effects can occur at more basic perceptual levels. For example, in the case of the well-known mere exposure effect preference increases with repeated exposures (Janiszewski, 1993; Monahan, Murphy & Zajonc, 2000), and has been shown across a variety of situations, including the subliminal level. Indeed, the mere exposure

Thomas Z. Ramsøy, Morten Friis-Olivarius, and Martin Skov, Decision Neuroscience Research Group, Department of Marketing, Copenhagen Business School, Frederiksberg, Denmark and Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark; Catrine Jacobsen and Simon B. Jensen, Decision Neuroscience Research Group, Department of Marketing, Copenhagen Business School.

AQ: 16

Correspondence concerning this article should be addressed to Thomas Z. Ramsøy, Decision Neuroscience Research Group, Department of Marketing, Copenhagen Business School, Solbjerg Plads 3, 2000, Frederiksberg, Denmark. E-mail: tzr.marktg@cbs.dk

effect demonstrates that preferences can be affected by purely perceptual mechanisms.

Little is still known about the fundamental psychological mechanisms underlying the effects that ambiguity and uncertainty has on preference, judgment, and choice. However, with the rise of modern cognitive neuroscience new possibilities have emerged for assessing the immediate perceptual, cognitive, and emotional mechanisms by which preference can be affected. Thus, studies have demonstrated that visual processing and emotion are related in a way such that perceptual features can influence the neural activity of brain structures thought to underlie emotional processing (e.g., Bechara, Damasio & Damasio, 2000; Delgado et al., 2008). In the same way, emotional responses to objects have been shown to modulate activity in the visual system. For example, the subjective valence of an object affects how attentional resources are allocated to perceptual processing of it. Thus, compared with neutral stimuli, more attentional resources are allocated to the perceptual computation of emotional, especially negative, objects (Öhman & Mineka, 2001; Algom, Chajut, & Lev, 2004; Estes & Verges, 2008), as demonstrated by an increase in detection accuracy (Nasrallah, Carmel, & Lavie, 2009).

As noted, the mere exposure effect shows that repeated exposure to a stimulus increases the preference for it compared with how well the same object is liked as a previous unfamiliar stimulus (Zajonc, 1968). In other words, a mere difference in perceptual exposure to an object is enough to enhance its perceived value, suggesting that the valuation process underlying the generation of preferences for experienced objects is not only susceptible to the content of what is sensed, but also to the sensory manner in which they are experienced. It remains unclear, however, exactly how changes in perception, as witnessed by the mere exposure effect, influence valuation processing, but one possible explanation, already put forward by Zajonc (1980, 2001), suggests that early perceptual computation signals affective structures to inform the brain quickly and automatically of possible biologically relevant changes to the environment. For example, unexpected changes could be signs of the sudden appearance of dangerous predators that would require swift action. The functional advantage of a perceptual system being able to provoke a fast affective response

would be quick changes in motivation for behavior without the brain having to cognitively represent and recognize what object it is in fact dealing with (Changizi, & Shimojo, 2008). Hence, it makes sense for unknown, novel, and unpredictable perceptual stimuli to be somewhat negatively colored by the brain's valuation system. As proposed by Zajonc, the reason why further exposure to an object leads to increased liking can be explained by the affective system signaling that the object turned out not to be dangerous after all. This explanation is supported by evidence that the mere exposure effect is larger when the stimulus is perceived without awareness (Bornstein, 1989; Bornstein, & D'Agostino, 1992), and that schizophrenia patients that exhibit impaired performance on explicit memory tasks show a normal mere exposure effect (Marie et al., 2001), suggesting that the impact of exposure on preference formation is attributable to unconscious, nonrepresentational processing. The latter point is further buttressed by studies showing that subliminal, affective primes can influence consumptive behavior, even when subjects experience no changes in conscious feelings (Winkielman, Berridge & Wilbarger, 2005). Indeed, although studies have focused on the possible role of subliminal cues in motivating subjects to approach goods and products (Veltkamp, Martijn & Aarts, 2011), and their limitations (Verwijmeren et al., 2011), little is known about the potential adverse effects that subliminal cues can have on preference and choice.

Support for this scenario can be found in recent neuroscience research that demonstrates that the perceptual systems of humans and animals are sensitive to alterations in the physical features of the surrounding sensory world, and that their affective systems are modulated by the sensory probability of an perceptual event (Kagan, 2009). Faced with some event that alters the immediate sensory surround, human brains will produce emotional reactions reflecting this unexpectedness. Specifically, brain structures important for generating negative emotions and avoidance behavior respond more strongly to novel relative to familiar information (Kiehl et al., 2001; Petrides, 2007; Wright et al., 2003). However, after repeated exposures such responses tend to decline (Yamaguchi, Hale, D'Esposito, & Knight, 2004).

A key brain structure thought to mediate the interaction between perception and emotion is the amygdala (Winkielman, Knutson, Paulus, & Trujillo, 2007; Pessoa & Adolphs, 2010). Amygdala activity has been demonstrated to predict preference for low-level perceptual properties, including an object's contour, in such a way that higher activity appears to reflect lower preference (Bar & Neta, 2007). A possible explanation for this phenomenon is that the amygdala reflects the salience or biological relevance of a perceptual event. Support for this hypothesis comes from studies that show that amygdala activity reflects a difference of impact of negative emotional images matched for valence and arousal (Ewbank et al., 2009), or from studies showing that just listening to an unexpected chord in a musical sequence elicits higher amygdala activation and lower preference (Koelsch, Fritz & Schlaug, 2008). In other words, the impact of amygdala activity on preference formation may be explained, in part at least, by its sensitivity to environmental contingencies in such a way that salient events, including previously unfamiliar stimuli, provokes a negative affective response which might be modulated in positive manner as the salience of the event changes, for instance through exposure.

Recent neuroimaging research has revealed that, indeed, the neural structure amygdala plays a pivotal role in computing emotional responses to perceptual expectedness. Several lines of evidence suggest that amygdala activity is sensitive to environmental contingencies and that changes in amygdala activity influences preference formation (e.g., Fischer et al., 2003; Ramsøy & Skov, 2010; Schwartz et al., 2003; Wright et al., 2003). Salient events, including unexpected stimuli, provoke a negative affective response, which might be modulated in positive manner as the salience of the event changes, for instance through repeated exposure to the object being evaluated. The link between novelty processing and emotional responses also extends to information that is ambiguous (Brand et al., 2007; Hsu et al., 2005; Schultz et al., 2008) or signals risk (Brand et al., 2007; Dreher, 2007). This implies that phenomena such as the mere exposure effect may be driven by a waning of novelty-based avoidance responses that is mediated through brain structures like the amygdala, because such structures

are involved in both novelty processing and negative emotions.

A recent study by Herry and colleagues (2007) directly supports this proposal. By playing neutral sound pulses in either an expected or unexpected sequence—that is, with either an predictable or unpredictable duration between the sounds—this seemingly biologically irrelevant feature engaged the amygdala. Besides suggesting that the amygdala may be involved in responding to environmental contingencies, this finding also indicates that information presented concurrently with an unexpected sound event may be perceived relatively less favorably. In the same study, playing predictable and unpredictable sound sequences concurrently with behavioral tasks led both mice and humans to respond in a more anxiety-like fashion, indicating that the response of the amygdala to perceptual unpredictability induced avoidance prone behavior. Herry and colleagues' data suggest that the response of the amygdala to unexpected environmental contingencies is critical to producing the kind of initial avoidance reaction to novel perceptual events we encounter in such cases as the mere exposure effect. Specifically, the waning of the initial aversive response to an unexpected event observed after repeated exposures (Zajonc, 1968; Zizak & Reber, 2004) is probably attributable to amygdalar habituation.

General Hypothesis

Based on these previous findings and theories, it is reasonable to assume that unpredictable factors can affect emotional responses and judgments. In particular, it is possible that such influences can work across different sensory domains, such as audition and vision. Based on this assumption, we hypothesize that contextual unpredictability is associated with lower preference for simultaneously presented information (H1).

To further qualify this effect, we suggest three additional hypotheses. If, as assumed, contextual unpredictability affects emotional responses and preferences, this effect should exist across different domains. Thus, we propose that contextual unpredictability will be associated with lower preference across different visual domains (H1a). Furthermore, although the study by Herry et al. (2007) demonstrated that

specific emotional brain structures, such as the amygdala, were more engaged by unpredictable sounds, this finding was not related to changes in physiological arousal and in itself provides little insight into the relationship between physiological arousal and emotional judgment. Based on this, we suggest two additional hypotheses: contextual unpredictability is related to increased physiological arousal (H1b), and that physiological arousal, when affected by contextual unpredictability, is negatively related to emotional judgment (H1c).

To test these hypotheses, we conducted three experiments. The first two studies tested the effect of contextual predictability—by using predictable and unpredictable sounds—on emotional preference for novel cultural objects such as brand logos and abstract art paintings, respectively. The third study used high-resolution eye-tracking to test the effect of contextual predictability on physiological arousal and its effect on subsequent preference ratings.

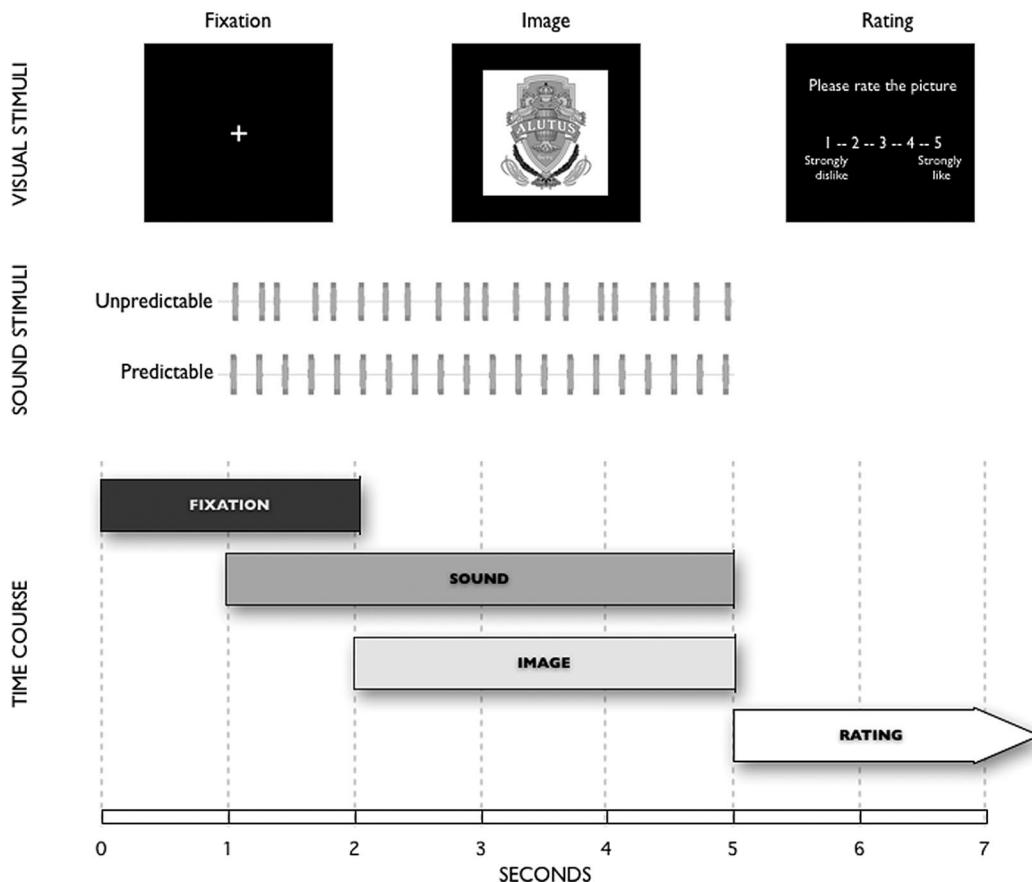
Study 1: Sound Ambiguity and Preference for Novel Brand Logos

If the general hypothesis is correct, one should observe preferences for objects vary when evaluated in situations where perceptual expectedness vary. The experimental paradigm developed by [Herry and his colleagues \(2007\)](#) unequivocally demonstrated that an unexpected sound sequence produced increased activation of the amygdala compared with an expected sound sequence, and that this neural effect was accompanied by elevated avoidance behavior both in humans and mice. If it is this neural mechanism that explains the mere exposure effect, we should observe higher preference ratings for novel brands when they are evaluated during perceptual predictability than when evaluated during perceptual unpredictability. To test this prediction we randomly assigned subjects to a two (sound: unpredictable or predictable) by four (brand logo category: cosmetics, electronics, finance, beer) factorial design study. Using the set-up from the [Herry et al. \(2007\)](#) study we presented unknown brand logos with either unpredictable or predictable sounds, and hypothesized that the logos perceived under the first condition would receive lower preference ratings than logos perceived under the latter condition.

Sixty subjects (38 women, age mean/std = 25.4/2.9, 52 right-handed) underwent experimental testing. All subjects were exposed to a total of 60 images of previously unknown brand logos, selected from a database of unused and unknown brand logos (www.brandsoftheworld.com). An equal number of brand logos from four different categories were used: cosmetics, beer, electronics, and finance. All stimuli were presented on a computer screen running E-Prime 2.0 (www.pstnet.com) in a Microsoft Windows XP environment. Visual stimuli were presented on a CRT monitor (resolution = 800 × 600 pixels; refresh rate = 85 Hz; average viewing distance = 60 cm). Sounds were presented through headphones (Sennheiser HD 555). Subjects responded using the numeric buttons on a computer keyboard.

For each trial, subjects first saw a white fixation cross on a black background. After 1 second, a predictable/unpredictable sound was played for 4 seconds. One second into the sound presentation, subjects saw a brand logo on the screen for 3 seconds. Thus, there was an overlap between sound stimulus and brand logo image for 3 seconds (see [Figure 1](#)). Among the total of 60 logo images, half (30) were randomly paired with the single predictable sound, and each of the other half was paired with one of the 30 unpredictable sounds. Following this, subjects were asked to rate their liking for each brand logo, using an on-screen five-point Likert scale (range: 1 = *strongly dislike*, 5 = *strongly like*). Response times were self-paced, that is, the test would only continue after the subject had responded. Responses and RTs were logged using E-Prime 2.

We constructed one predictable and 30 unpredictable sounds with a duration of 4 seconds, following the method outlined in [Herry et al. \(2007\)](#). Predictable sounds were made with a carrier frequency of 1 kHz with pulse duration of 40 ms and a pulse spacing of 200 ms (5 Hz pulse repetition rate). Unpredictable sound pulse sequences were constructed by applying a random jitter to the timing of the sound pulses, with a maximum offset of ± 60 ms. The sound pulse time series was derived by randomly varying pulse timing within defined temporal boundaries of predictable sound pulse sequences with regular pulse timing. The sound pulse was switched on and off smoothly using



AQ: 17-18

Figure 1. Overview of study design. Subjects first saw a fixation cross, which was accompanied by a sound after 1 second. At second 3 the fixation cross was replaced by a brand logo. At second 5 subjects were shown a Likert scale and asked to rate the previously seen brand logo.

F2

cosine-shaped increasing and decreasing ramps with a duration of 5 ms.

The statistical analysis first focused on the *a priori* hypothesis, by using a linear mixed effect (variance component) model. The dependent variable was liking response, with sound type (predictable, unpredictable) as independent variable, using subject as a random factor. To test for effects of sound on reaction time (RT), we also used the same linear mixed model analysis with RT as a dependent variable. Responses with a lower response time than 150 ms were excluded. Because we used a directional hypothesis, we performed a one-tailed test of significance.

After this analysis, we performed exploratory analyses by including additional variables and

their interactions. Here, we used a linear mixed model, with gender (male, female), sound type (predictable, unpredictable), brand logo category (electronic, finance, cosmetics, beer; see Figure 2), and with subject as a random factor. We used a stepwise backward elimination procedure to obtain an optimal explanatory model of liking ratings. All analyses were run in R version 2.1.0 (<http://cran.r-project.org/>).

Results

AQ: 3

The main effect of sound compared the relationship between sound type with preference responses. The mixed effects analysis confirmed the directional hypothesis: brand logos accompanied by unpredictable sounds were

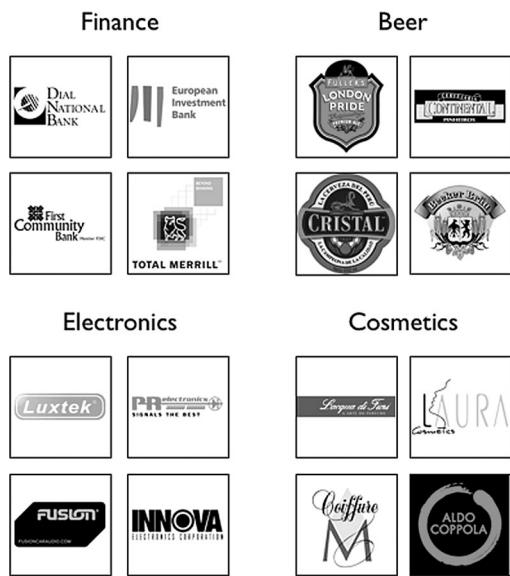


Figure 2. Examples of brand logos used in the studies. Brand logos belonged to either of four categories, including finance (top left), beer (top right), electronics (bottom left), and cosmetics (bottom right). No brand logo was previously known to the subjects.

rated significantly lower (mean/std = 2.37/0.07) than logos accompanied by predictable sounds (mean/std = 2.43/0.07; $t = 1.72, p = .043$). The effects are also displayed in Figure 3. There was no significant difference in RT between the sound types (unpredictable = 0.146 ± 0.002 ms; predictable = 0.146 ± 0.002 ms, $t = 0.049, p = .961$), indicating that there were no differences in processing demand between the two sound types.

In our exploratory mixed model analysis we initially included variables (sound type, brand type, gender, and with subject as random factor) and their interactions and used a stepwise backward elimination procedure. The final model included image type, gender, and the gender by image interaction effect (see Table 1).

The observed gender effect, taking sound effects into account, was driven by a more positive rating by men (mean/std = 2.640/0.116) compared with women (mean/std = 2.256/0.088, $p = .011$). Performing pairwise comparisons of brand logo categories, as Table 2 illustrates, revealed that beer logos received the highest ratings, followed by cosmetics, bank, and finally electronics logos. An analysis of the ef-

fects of brand category on RT showed no significant effect, suggesting that the effects of brand category were not the result of increased cognitive processing demand.

Finally, the study of interaction effects between gender and brand logo category revealed that besides the main effect higher ratings in men on all logos, the difference in preference for the two genders were particularly pronounced for beer and cosmetics logos.

Discussion

The aim of Study 1 was to test the extent to which contextual unpredictability could affect preference for simultaneously presented information. In particular, we were interested in testing whether unpredictability within one domain (audition) could affect emotional processing of information in a different domain (vision). In this first study, we demonstrate that the hedonic judgment of unknown brand logos is affected by a simultaneously played sound. If the sound was unpredictable, subjects' judgments of brand logos were significantly lower than if the accompanying sound was predictable. This finding is in line with the recent study by [Herry et al. \(2007\)](#), in which the perception of unpredictable sounds, which led to an increased engagement of the amygdala, was associated with avoidance behaviors. Taken together with our data, in which we used identical sound stimuli, this suggests that contextual unpredictability engages the brain's aversion circuit, and that this effect affects the subsequent emotional processing of simultaneously occurring information. Our results are in line with studies demonstrating adverse effects of contextual variables such as auditory and visual noise on cognitive performance and emotions ([Nagar & Pandey, 1987](#); [Haines et al., 2001](#); [Schupp et al., 2008](#)). For example, in a study by [Keizer, Lindenberg, and Steg \(2008\)](#) it was shown that environmental noise and disorder would affect people to act more disorderly and immoral. Our results provide a more detailed insight into these effects by suggesting that the immediate effects of contextual noise is an increase in aversive emotional responses that affect simultaneously presented information.

Notably, the results in this first study demonstrate that the effects of contextual predictability on preference is limited, and that it is overshadowed

F3

AQ: 4

T1

T2

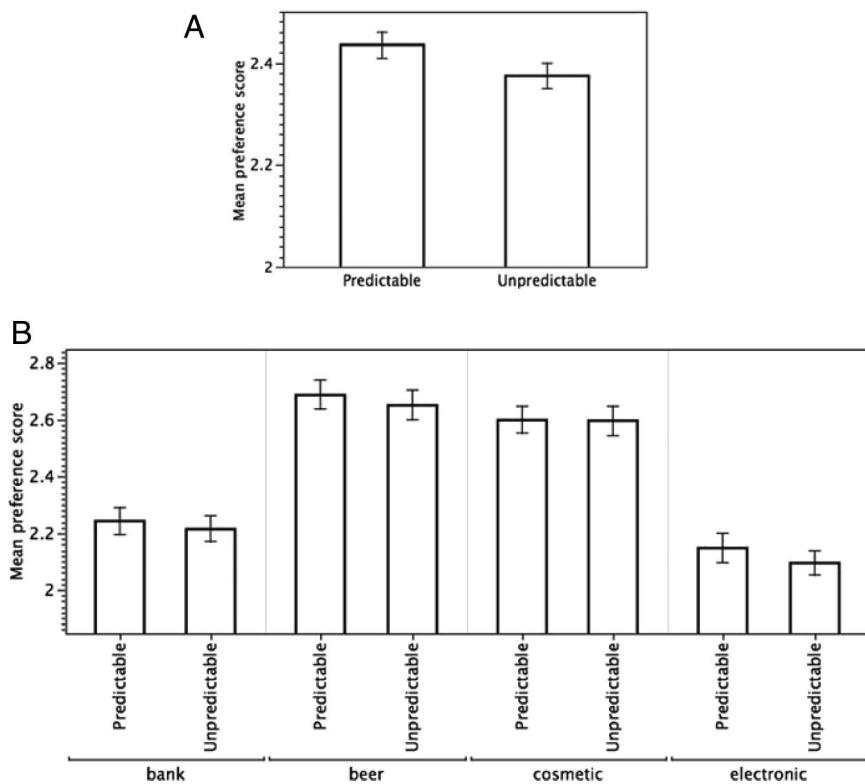


Figure 3. Effects of sound type on preference for novel brand logos. The main effect (A) of sound type shows a significant lower preference score for logos coupled to unpredictable sounds compared with predictable sounds. For each brand category (B), this effect is overshadowed by the large inter-category difference in preference scores. In both plots, bars indicate mean preference score, whiskers indicate standard error of the mean.

owed by other, larger effects such as logo type and gender. Nevertheless, although such effects were small, the use of a seven-point Likert scale may also have limited the sensitivity of the scale for detecting subtle changes in preference. Furthermore, our study only used sound as the carrier for stimulus predictability, and thus fur-

ther research would be needed to test the effect in other sensory domains.

It should also be noted that our observations are relative differences between predictable and unpredictable sounds. Therefore, the data should be interpreted with some caution, as the results from Study 1 do not provide information about the directionality of the effects. However, as we shall see, when we explore the emotional responses to the manipulations (Study 3), we find additional evidence for directionality.

The observed effects in Study 1 may be the result of a unique combination of contextual cues and visual information. To better understand the generalization of the effects of contextual unpredictability on emotional processing, we wanted to test whether other kinds of visual information would be affected similarly to novel brand logos. To this end, we con-

Table 1
Effects of Category, Gender, and Interaction on Preference Judgments

Variable	df	F	p
Category	(3;3366.231)	79.397	<0.001
Gender	(1;57.977)	6.971	0.011
Category * Gender	(3;366.231)	5.119	0.002

Note. Results from exploratory analysis using backward stepwise regression, demonstrating main effects of brand category, gender, and their interaction.

Table 2
Effects of Brand Category on Preference Judgments

Brand category (I)	Brand category (J)	Mean difference* (I – J)	Std. error	Sig.	95% confidence interval for difference	
					Lower bound	Upper bound
Bank	Beer	-.455*	0.042	0.000	-0.537	-0.372
	Cosmetic	-.345*	0.042	0.000	-0.428	-0.263
	Electronic	.095*	0.042	0.025	0.012	0.177
Beer	Bank	.455*	0.042	0.000	0.372	0.537
	Cosmetic	.109*	0.042	0.009	0.027	0.191
	Electronic	.549*	0.042	0.000	0.467	0.631
Cosmetic	Bank	.345*	0.042	0.000	0.263	0.428
	Beer	-.109*	0.042	0.009	-0.191	-0.027
	Electronic	.440*	0.042	0.000	0.358	0.522
Electronic	Bank	-.095*	0.042	0.025	-0.177	-0.012
	Beer	-.549*	0.042	0.000	-0.631	-0.467
	Cosmetic	-.440*	0.042	0.000	-0.522	-0.358

* Differences in preference ratings between brand categories, when taking sound effects into account.

structed a second study, in which we tested the effect of sound predictability on the emotional judgments of abstract art.

Study 2 – Sound Ambiguity and Judgment of Art

The experimental set-up was the same as in Study 1. The same 60 subjects were used, but instead of 60 images of brand logos in this study we used 60 abstract paintings, taken from various Internet resources. These 60 paintings were selected from a larger pilot session of 200 images, which were rated for aesthetic liking by a separate group of 10 subjects. The 60 images with middle ratings across these subjects were chosen. As in Study 1, the paintings were presented for 3 seconds, followed by a self-paced rating task, in which subjects used a Likert scale to report their subjective liking of the image (range 1–5, where 1 = *strongly dislike*, 5 = *strongly like*). Sound stimuli were presented 1 second before image onset, with a duration of 4 seconds.

As with Study 1, we analyzed the data using a linear mixed model with liking as the dependent variable, and with sound type as the independent variable and subject as random factor. The analysis was run using R version 2.1.0.

Results

Rating during unpredictable sounds (mean \pm std = 2.84 ± 0.03) were significantly lower

than ratings during predictable sounds (2.96 ± 0.03), $t = 3.04$, $df = (1; 3362.5)$, $p = .0012$, and the effects are displayed in Figure 4. In an explorative analysis we included subject as random factor, to abate the effects of individual preferences for art. In doing so, we find an even stronger effect ($t = 3.4$, $p = .0007$), and with the model explaining a quarter of all the variation in the ratings ($R^2 = 0.252$). There was no effect of sound type on response time (predictable = 850 ± 21 ms; unpredictable = 856 ± 22 ms, $t = 0.21$, $p = .835$).

Discussion

The aim of Study 2 was to test to what extent the effect of contextual contingency on prefer-

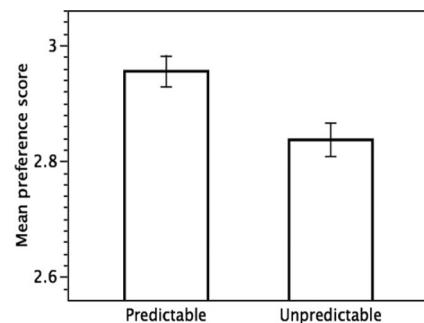


Figure 4. Effect of sound type on preference for abstract art. When images of art were coupled to unpredictable sounds they were rated significantly lower than when they were coupled to a predictable sound. Graph displays mean preference score, whiskers denote standard error of the mean.

ence for brand logos would work on a different visual domain, such as abstract art. The study conclusively demonstrates that sound predictability also shows similar effects on preference for other visual elements. Indeed, compared with the relative small effects seen in Study 1, sound predictability had a robust effect on preference for abstract art. While caution should be exercised for such findings in which there is no direct comparison between the two conditions, one possible explanation could be that the two conditions differed on at least two parameters. On the one hand, the abstract art paintings could have a wider range of layouts, as opposed to a relatively more constrained visual appearance seen in brands. This would induce more variability in ratings of art, and thus a larger space from which effects of sound predictability could be detected. Second, one may contend that at a more social—and even societal—level, judgment of abstract art may be seen as a more individualistic assertion than brand logos. This could induce a stronger variability in ratings for art images, thus producing a broader range of preference data for detecting changes induced by sound predictability.

Despite using identical stimuli to that of [Henry et al. \(2007\)](#), in which there was a significant engagement of the amygdala, the use of such stimuli under other conditions could still lead to differences in emotional processing. To test whether the use of unpredictable sounds lead to emotional arousal in our experimental conditions, we therefore ran a third study in which we assessed subjects' physiological arousal to predictable and unpredictable sound stimuli, and subsequent sound plus visual combinations. Besides demonstrating that our use of contextual predictability worked in our experimental set-up, we could also learn more about the physiological bases of the effects we observed.

Study 3 – Effects of Contextual Ambiguity on Arousal and Preference Judgments

In our third and final study, we wanted to test to what extent the effect of sound unpredictability on preference would be associated with changes in arousal. If, as shown by [Henry et al. \(2007\)](#), unpredictable sounds are associated with increased amygdala activation and avoidance behavior, it stands to reason that other

measures of arousal will show similar deflections. It is well established in the neuroscientific and neurophysiological literature that the amygdala is involved in the regulation of bodily responses such as facial expressions, pulse, respiration, sweating, and pupil dilation ([De Martino, Camerer, & Adolphs, 2010](#); [Groepel-Klein, 2005](#); [Steinmetz, Addis & Kensinger, 2010](#)). As such responses tend to be collinear ([Bradley et al., 2008](#)), using one measure would suffice as an index of arousal.

By using high-resolution eye-tracking we sought to test whether unpredictable sounds would lead to a higher arousal responses, as indexed by pupil dilation, and thus test H1b. If, as assumed, unpredictable sounds are associated with negative emotional responses, this would also lead to stronger pupillary responses, when compared with predictable sounds. Furthermore, the use of pupil dilation in our experimental set-up would also serve as a validation of the main method in demonstrating the involvement of arousal as a main factor in our behavioral effects.

Method

Eighteen right-handed subjects (age 24.8 ± 5.3 years, eight women) with normal or corrected to normal vision were recruited from the Copenhagen region. Based on a self-report questionnaire, no subject experienced, or had any indication of, neurological or psychiatric disease. All subjects provided informed consent.

High-resolution eye-tracking was performed using a Tobii T60 XL tracker running at 60 Hz with a 1920×1200 pixel screen resolution and an approximate viewing angle of 60 cm. Stimulus presentation and the recording of subject responses and eye tracking data were performed using Attention Tool version 4.5 (iMotions Inc., www.imotionsglobal.com).

After undergoing a nine-point eye tracking calibration procedure, subjects performed the same experimental task as in the two prior studies; after a first second of exposure to a predictable or unpredictable sound, a novel brand logo was shown for 3 seconds while the same sound continued. The sound stimuli were identical to the prior studies, and used pseudorandomly with brand logos, and in which sound-image coupling was counterbalanced across subjects.

As in the prior study using brand logos, we used four categories of brand logos (finance, cosmetics, beer, and electronics). Subjects rated their preference for the brand logo using an analogue scale on screen using a computer mouse.

All data were analyzed using JMP version 9.0 (SAS Inc.). We first analyzed the effects of sound type on preference ratings by running a general linear model (GLM) analysis with ratings as the dependent variable, using sound type (predictable, unpredictable) as the independent variable, and with subject as random factor. To analyze the general effect of sound type on pupil dilation we then ran a second GLM analysis using pupil dilation as the dependent variable and with sound type as the independent variable. A follow-up analysis looked at whether this effect was different between the pure sound condition and the simultaneous sound and brand logo presentation. To analyze this we ran a GLM analysis in which pupil dilation was used as the dependent variable, and with sound type (predictable, unpredictable), slide type (sound, sound + logo) and their interaction was used as independent variables, and with subject as random factor. Finally, we ran an explorative analysis of the effects of pupil dilation on preference ratings using preference rating as the dependent variable, and with pupil dilation as the independent variable.

Results

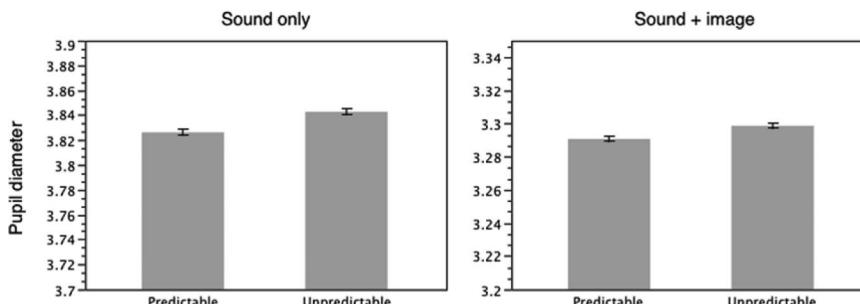
Unpredictable sounds were associated with lower preference ratings ($\text{mean} \pm \text{SEM} = 89.18 \pm 0.09$) when compared with predictable

sounds (89.41 ± 0.1 , $t = 1.76$, $p = .039$ one-tailed). Running an explorative regression analysis with preference score as the dependent variable and with sound type, brand category and sound type*brand category interaction, and with subject as random factor, we find a main effect of brand category ($F = 2039.2$, $p < .0001$) and a sound type*brand category interaction ($F = 190.4$, $p < .0001$), but no longer any main effect of sound type ($F = 0.2$, $p = .643$).

Our first analysis of the main effect of sound type on pupil dilation demonstrated a significant effect where unpredictable sounds led to a significantly stronger pupil dilation response ($\text{mean} \pm \text{SEM} = 3.44 \pm 0.007$) than predictable sounds did (3.42 ± 0.007 , $t = 7.7$, $p < .0001$), as shown in [Figure 5](#).

A follow-up analysis looked at whether the sound effect was different for the two stimulus conditions, where either the sound was presented alone, or where sound and logo was presented simultaneously. Here we found that besides the effect of sound type ($F_{\text{predictable, unpredictable}} = 52.9$, $t = -7.3$, $p < .0001$) there was an additional strong effect of stimulus condition ($F_{\text{sound, sound + image}} = 91348.0$, $t = 302.3$, $p < .0001$) but no interaction effect ($F = 2.0$, $t = -1.41$, $p = .160$; see also [Figure 6](#)). This effect is most likely caused by a pupil restriction to the increased brightness due to brand logo presentation.

Our final analysis focused on the relationship between pupil dilation and preference, by using preference ratings as the dependent variable,



[Figure 5](#). Pupil dilation effects of predictable and unpredictable sounds. In the condition with sound only (left) unpredictable sounds produce significantly stronger pupil dilation than predictable sounds. In the combined sound and image condition (right) this effect is still present, albeit somewhat weaker. Bars represent mean pupil diameter, whiskers denote standard error of the mean.

F5

F6

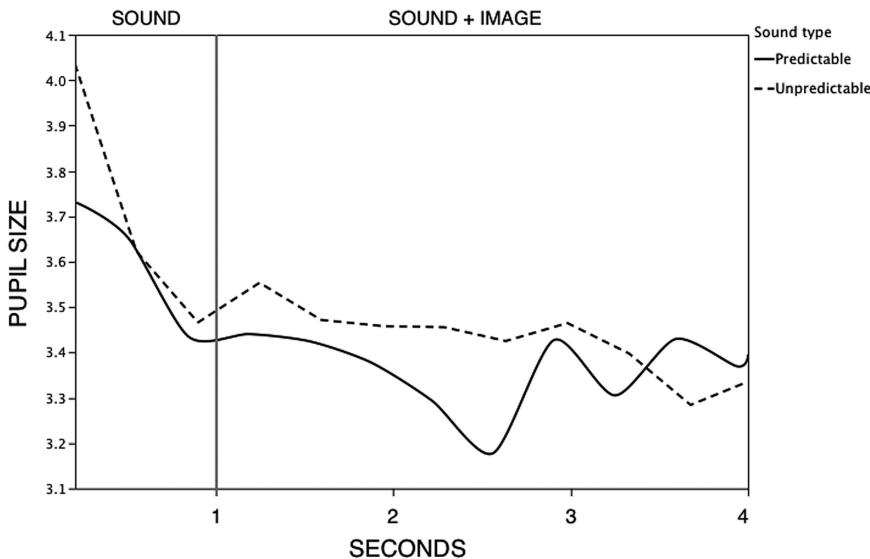


Figure 6. Effects of sounds on pupil dilation over time. During the first second of pure sound, there is an initial pupil dilation response that is significantly stronger for unpredictable sounds (dotted line) than predictable sounds (solid line). When sound and image was presented simultaneously, unpredictable sounds still produced a stronger average pupil dilation, especially during the first half of this period. Please note that the drop in pupil diameter before the onset of the image is most likely a conditioned response to the consistent presentation of the bright image after 1 second of pure sound and dark screen. Each line indicates mean pupil diameter.

pupil dilation the independent variable, and with subject as random factor in a GLM analysis, one analysis for the pure sound condition and one for the sound plus brand logo.

For the pure sound condition we find a negative linear relationship between pupil dilation and preference ($R^2 = 0.29$, $F = 40.6$, $\beta = -0.05$, $p < .0001$), suggesting that stronger pupil dilation was related to subsequent lower preference scores. For the condition with sound and logo, we find a positive linear effect ($R^2 = 0.19$, $F = 98.2$, $\beta = 0.02$, $p < .0001$). To further qualify this, we ran an additional post hoc analysis of the interaction between pupil size and epoch, and found this to be highly significant ($F = 17.7$, $p < .0001$).

Discussion

Two issues were sought to be resolved with Study 3. First, our choice of sounds closely mimicked the procedure used by [Herry et al. \(2007\)](#), in which unpredictable sounds evoked stronger emotional responses (assessed by the engagement of the amygdala, and the extent of

avoidance behavior) compared with predictable sounds. Nevertheless, as our study used a different kind of stimuli and emotional responses (preference ratings), it was important to demonstrate that unpredictable sounds did evoke stronger emotional arousal than predictable sounds. Indeed, this was found, as unpredictable sounds produced larger pupil dilation responses compared with predictable sounds. Moreover, we were able to establish that this effect was present both when the pure tone was presented, and in the condition where subjects saw a brand logo and heard the sound. This finding suggests that an unpredictable sound evokes emotional arousal when presented alone, and that this effect spills over to the emotional processing of subsequently presented visual information. We believe that this arousal effect is responsible for the difference in brand logo preference, that is, that unpredictable sounds evoke an arousal response which leads to an initial negative emotional evaluation of simultaneously presented information in other visual domains.

It should be noted that Study 3 provides more information about the directionality of our behavioral results. While, on the one hand, both predictable and unpredictable sounds produced positive deflections in pupil dilation (i.e., an emotion response), we found that these responses were significantly stronger in the unpredictable condition. This finding suggests that the relative difference in preference, as seen in Study 1 and 2, is associated with a relatively stronger emotional response to unpredictable sounds.

Our second aim was to test the relationship between pupil dilation and preference. By running a regression analysis with pupil dilation as the independent variable and preference as the dependent variable, and using subject as a random factor, we find that the relationship is different depending on whether the sound is presented alone or together with the brand logo. In the condition with the single sound, we find a negative relationship between pupil size and brand logo preference. That is, the stronger the pupil response during this phase, the lower the subsequent brand logo preference. When analyzing the second condition, where sound and brand logo are presented simultaneously, we find an opposite effect. Here, larger pupil size in this period is related to more positive preference ratings. Our follow-up analysis furthermore demonstrated that this interaction effect between pupil dilation and stimulus condition was significant. This finding suggests that the negative effect of sound type on preference is most likely caused by the arousal response during the pure sound condition, and that, if anything, pupil dilation has no adverse effect on preference ratings when sounds are presented simultaneously with brand logos.

General Discussion

Judgments and choices have long been known to be affected by contextual factors such as uncertainty and ambiguity, and that even perceptual properties such as visual salience and mere repetition can lead to alterations in preference and judgments. However, our understanding of the exact mechanisms underlying these effects, and to what extent they can be induced through contextual stimuli alone, is still limited. The aim of this study was to test to what extent preferences would be influenced by

simple contextual factors such as unpredictable sounds, and to what extent emotional responses could account for such effects. Based on recent studies that report increased aversion-related brain activation and avoidance behavior by rodents and humans to items associated with contextual unpredictability, we set out to test whether a similar set-up could affect preferences for cultural items. In particular, we tested whether sound predictability could operate as a key factor influencing preference for simultaneously presented complex visual items, such as novel brand logos and abstract art. Here, we report that preferences for visual stimuli belonging to different categories were negatively affected if items were simultaneously presented with unpredictable sounds, as opposed to a predictable sound. Furthermore, by using high resolution eye-tracking, we link these effects of sound on preference judgments to immediate changes in arousal, as indexed by pupil dilation responses.

Our findings provide novel insights into the mechanisms that contextual ambiguity and uncertainty can have on judgments and behavior. Our data bring together two strings of research on preference and judgments. First, behavioral economics have demonstrated that uncertainty and ambiguity, as explicit parts of decision contexts, increase aversion, and avoidance behaviors. Recent inclusions of physiological and neurobiological approaches have extended these findings by relating such behavioral effects to increased arousal (e.g., Bechara, Damasio & Damasio, 2000) and the engagement of brain structures involved in fear and aversion (Christopoulos et al., 2009). Here, our findings demonstrate that simple perceptual effects, such as the temporal ordering of meaningless sounds, can produce similar effects on arousal and subsequent judgments. We contend that these findings should be tested using stimulus predictability in other sensory domains, such as visual, tactile, and olfactory stimuli, when possible. Second, studies on the psychology of preference formation have long demonstrated different shades of the so-called mere exposure effect, in which stimulus repetition alone produces an increase in preference judgments (or, alternatively, a reduction of initial skepticism toward a novel object). Although our studies did not use repeated stimuli, they may still inform our understanding of the basic mechanisms of mere

exposure effects. As noted by Zajonc (1980, 2001) one explanation of the mere exposure effect is that the positive relationship between the number of presentations and preference may reflect a waning aversion, rather than a waxing of positive emotions. In this study, our demonstration that perceptual unpredictability leads to increased arousal and lower preference judgments could suggest that first impressions bear on an initial aversion response.

This still leaves out questions that remain to be answered. First, it is yet unknown whether simple contextual cues such as unpredictable sounds can affect value-based decision-making at higher levels such as gambling, multiattribute choices, and social dilemmas. Although these effects and limitations need to be studied better, it is still expected that any effect of ambiguity, as used here, will rely on the same basic emotional responses. To what extent emotional responses are induced by contextual ambiguity in all contexts, but are controlled by higher-order cognitive functions in certain situations but not others, remains to be seen.

AQ: 5

References

- Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: A generic slowdown, not a stroop effect. *Journal of Experimental Psychology: General*, 133, 323.
- Bar, M., & Neta, M. (2007). Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia*, 45, 2191–2200. doi:10.1016/j.neuropsychologia.2007.03.008
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10, 295–307. doi:10.1093/cercor/10.3.295
- Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, 63, 545–552. doi:10.1037/0022-3514.63.4.545
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45, 602–607. doi:10.1111/j.1469-8986.2008.00654.x
- Brand, M., Grabenhorst, F., Starcke, K., Vandekerckhove, M. M. P., & Markowitsch, H. J. (2007). Role of the amygdala in decisions under ambiguity and decisions under risk: Evidence from patients with urbach-wiethe disease. *Neuropsychologia*, 45, 1305–1317. doi:10.1016/j.neuropsychologia.2006.09.021
- Changizi, M. A., & Shimojo, S. (2008). A functional explanation for the effects of visual exposure on preference. *Perception*, 37, 1510–1519. doi:10.1080/p6012
- Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *The Journal of Neuroscience*, 29, 12574–12583. doi:10.1523/JNEUROSCI.2614-09.2009
- Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, 29, 537–545. doi:10.1016/S0896-6273(01)00225-2
- Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, 363, 3787–3800. doi:10.1098/rstb.2008.0161
- De Martino, B., Camerer, C. F., & Adolphs, R. (2010). Amygdala damage eliminates monetary loss aversion. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 107, 3788–3792. doi:10.1073/pnas.0910230107
- Dreher, J. C. (2007). Sensitivity of the brain to loss aversion during risky gambles. *Trends in Cognitive Sciences*, 11, 270–272. doi:10.1016/j.tics.2007.05.006
- Estes, Z., & Verguts, M. (2008). Freeze or flee? Negative stimuli elicit selective responding. *Cognition*, 108, 557–565.
- Ewbank, M. P., Barnard, P. J., Croucher, C. J., Ramponi, C., & Calder, A. J. (2009). The amygdala response to images with impact. *Social Cognitive and Affective Neuroscience*, 4, 127–133. doi:10.1093/scan/nsn048
- Fischer, H., Wright, C. I., Whalen, P. J., McInerney, S. C., Shin, L. M., & Rauch, S. L. (2003). Brain habituation during repeated exposure to fearful and neutral faces: A functional MRI study. *Brain Research Bulletin*, 59, 387–392. doi:10.1016/S0361-9230(02)00940-1
- Fukui, H., Murai, T., Fukuyama, H., Hayashi, T., & Hanakawa, T. (2005). Functional activity related to risk anticipation during performance of the iowa gambling task. *NeuroImage*, 24, 253–259. doi:10.1016/j.neuroimage.2004.08.028
- Groepel-Klein, A. (2005). Arousal and consumer in-store behavior. *Brain Research Bulletin*, 67, 428–437. doi:10.1016/j.brainresbull.2005.06.012
- Haines, M. M., Stansfeld, S. A., Job, R. F., Berglund, B., & Hea, J. (2001). Chronic aircraft noise exposure, stress responses, mental health and cognitive performance in school children. *Psychological*

- Medicine*, 31, 265–277. doi:10.1017/S0033291701003282
- AQ: 6 Herry, C., Bach, D. R., Esposito, F., Di Salle, F., Perrig, W. J., Scheffler, K., . . . Seifritz, E. (2007). Processing of temporal unpredictability in human and animal amygdala. *The Journal of Neuroscience*, 27, 5958–5966. doi:10.1523/JNEUROSCI.5218-06.2007
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310, 1680–1683. doi:10.1126/science.1115327
- Janiszewski, C. (1993). Preattentive mere exposure effects. *Journal of Consumer Research*, 20, 376–392. doi:10.1086/209356
- Kagan, J. (2009). Categories of novelty and states of uncertainty. *Review of General Psychology*, 13(4), 290.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–292. doi:10.2307/1914185
- Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, 322(5908), 1681–1685. doi:10.1126/science.1161405
- Kensinger, E. A., & Schacter, D. L. (2008). Neural processes supporting young and older adults' emotional memories. *Journal of Cognitive Neuroscience*, 20, 1161–1173. doi:10.1162/jocn.2008.00080
- Kiehl, K. A., Laurens, K. R., Duty, T. L., Forster, B. B., & Liddle, P. F. (2001). Neural sources involved in auditory target detection and novelty processing: An event-related fMRI study. *Psychophysiology*, 38, 133–142. doi:10.1111/1469-8986.3810133
- Koelsch, S., Kilches, S., Steinbeis, N., & Schelinski, S. (2008). Effects of unexpected chords and of performer's expression on brain responses and electrodermal activity. *PloS ONE*, 3, e2631. doi:10.1371/journal.pone.0002631
- Levy, I., Snell, J., Nelson, A. J., Rustichini, A., & Glimcher, P. W. (2010). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology*, 103, 1036–1047. doi:10.1152/jn.00853.2009
- Marie, A., Gabrieli, J. D. E., Vaidya, C., Brown, B., Pratto, F., Zajonc, R. B., & Shaw, R. J. (2001). The mere exposure effect in patients with schizophrenia. *Schizophrenia Bulletin*, 27, 297.
- Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science: A Journal of the American Psychological Society*, 11, 462–466.
- Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393(6684), 467–70. doi:10.1038/30976
- Muthukrishnan, V., Wathieu, L., & Xu, J. (2009). Ambiguity aversion and the preference for established brands. *Management Science*, 55, 1933–1941. doi:10.1287/mnsc.1090.1087
- Nagar, D., & Pandey, J. (1987). Affect and performance on cognitive task as a function of crowding and noise. *Journal of Applied Social Psychology*, 17, 147–157.
- Nasrallah, M., Carmel, D., & Lavie, N. (2009). Murder, she wrote: Enhanced sensitivity to negative word valence. *Emotion*, 9, 609–618.
- Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483–522.
- Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, 11, 773–783.
- Petrides, M. (2007). The orbitofrontal cortex: Novelty, deviation from expectation, and memory. *Annals of the New York Academy of Sciences*, 1121, 33–53. doi:10.1196/annals.1401.035
- Ramsøy, T. Z., & Skov, M. (2010). How genes make up your mind: Individual biological differences and value-based decisions. *Journal of Economic Psychology*, 31, 818–831. doi:10.1016/j.jeop.2010.03.003
- Schultz, W., Preuschoff, K., Camerer, C., Hsu, M., Fiorillo, C. D., Tobler, P. N., & Bossaerts, P. (2008). Explicit neural signals reflecting reward uncertainty. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1511), 3801–3811. doi:10.1098/rstb.2008.0152
- Schupp, H. T., Stockburger, J., Schmälzle, R., Bublitzky, F., Weike, A. I., & Hamm, A. O. (2008). Visual noise effects on emotion perception: Brain potentials and stimulus identification. *NeuroReport*, 19, 167–171. doi:10.1097/WNR.0b013e3282f4aa42
- Schwartz, C. E., Wright, C. I., Shin, L. M., Kagan, J., Whalen, P. J., McMullin, K. G., & Rauch, S. L. (2003). Differential amygdalar response to novel versus newly familiar neutral faces: A functional MRI probe developed for studying inhibited temperament. *Biological Psychiatry*, 53, 854–862. doi:10.1016/S0006-3223(02)01906-6
- Smith, B. W., Mitchell, D. G., Hardin, M. G., Jazbec, S., Fridberg, D., Blair, R. J., & Ernst, M. (2009). Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *NeuroImage*, 44, 600–609. doi:10.1016/j.neuroimage.2008.08.016
- Steinmetz, M., K. R., Addis, D. R., & Kensinger, E. A. (2010). The effect of arousal on the emo-

AQ: 10

AQ: 11

AQ: 13

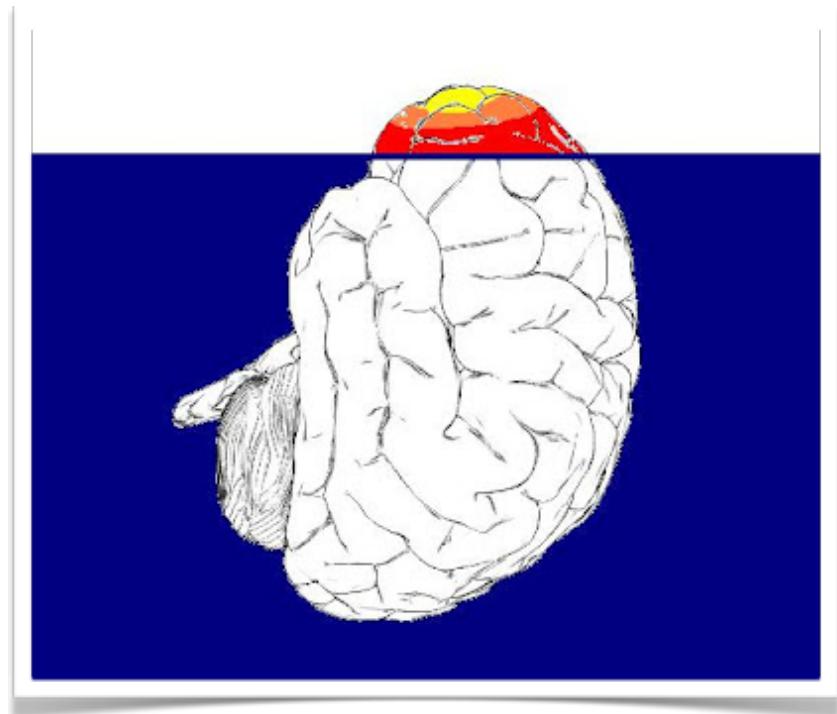
- tional memory network depends on valence. *NeuroImage*, 53, 318–324. doi:10.1016/j.neuroimage.2010.06.015
- Veltkamp, M., Custers, R., & Aarts, H. (2011). Motivating consumer behavior by subliminal conditioning in the absence of basic needs: Striking even while the iron is cold. *Journal of Consumer Psychology*, 21, 49–56. doi:10.1016/j.jcps.2010.09.011
- Verwijmeren, T., Karremans, J. C., Stroebe, W., & Wigboldus, D. H. J. (2011). The workings and limits of subliminal advertising: The role of habits. *Journal of Consumer Psychology*, 21, 206–213. doi:10.1016/j.jcps.2010.11.004
- Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Emotion, behavior, and conscious experience: Once more without feeling. In L. Feldman-Barrett, P. Niedenthal, & P. Winkielman (Eds.), *Emotion and consciousness*. New York: Guilford Press.
- Winkielman, P., Knutson, B., Paulus, M., & Trujillo, J. L. (2007). Affective influence on judgments and decisions: Moving towards core mechanisms. *Review of General Psychology*, 11, 179–192.
- Wright, C. I., Martis, B., Schwartz, C. E., Shin, L. M., Fischer, H. H., McMullin, K., & Rauch, S. L. (2003). Novelty responses and differential effects of order in the amygdala, substantia innominata, and inferior temporal cortex. *Neuroimage*, 18, 660–669. doi:10.1016/S1053-8119(02)00037-X
- Yamaguchi, S., Hale, L. A., D'Esposito, M., & Knight, R. T. (2004). Rapid prefrontal-hippocampal habituation to novel events. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 24, 5356–5363.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, Monograph supplement No. 2, Part 2.
- Zajonc, R. B. (1980). Feelings and thinking: Preferences need no inferences. *American Psychologist*, 35, 151–175. doi:10.1037/0003-066X.35.2.151
- Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10, 224–228. doi:10.1111/1467-8721.00154
- Zajonc, R. B., & Markus, H. (1982). Affective and cognitive factors in preferences. *Journal of Consumer Research*, 9, 123–131. doi:10.1086/208905
- Zizak, D. M., & Reber, A. S. (2004). Implicit preferences: The role(s) of familiarity in the structural mere exposure effect. *Consciousness and Cognition: An International Journal*, 13, 336–362. doi:10.1016/j.concog.2003.12.003

AQ: 14

AQ: 15

Received July 2, 2012
 Revision received August 21, 2012
 Accepted August 23, 2012 ■

ATTENTION, AWARENESS & CONSCIOUSNESS



One of the main reasons for the rise of neuromarketing and consumer neuroscience is the realization that human decision making deviates from “rational choice”. Despite our experience of the contrary, our decisions are not particularly rational or optimal; we do not make decisions after carefully exploring all options; and our decisions are rarely driven by conscious planning.

Rather, consumer choices are driven by at least two main processes: an unconscious “quick and dirty” valuation of immediate options, and conscious experiences of making the choice. The puzzling thing is that whenever we study human decision making, we find that brain activation predating our experience of actually making the choice can account for most of the actual choice.

To better understand the unconscious and conscious mind, we need to look at some of the basics. What is the conscious mind? How much of our behavior is actually driven by unconscious processes?

Also, attention is a mongrel concept that we need to tease apart. Whether your attention is driven by external factors or your conscious will is not trivial. And how much “bandwidth” do you think your brain has? Do you have any mental bottlenecks?

Brain, conscious experience and the observing self

Bernard J. Baars¹, Thomas Z. Ramsøy² and Steven Laureys³

¹The Neurosciences Institute, San Diego, CA 92121, USA

²Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital, DK-2650 Hvidovre, Denmark

³Cyclotron Research Centre, University of Liège, 4000 Liège, Belgium

Conscious perception, like the sight of a coffee cup, seems to involve the brain identifying a stimulus. But conscious input activates more brain regions than are needed to identify coffee cups and faces. It spreads beyond sensory cortex to frontoparietal association areas, which do not serve stimulus identification as such. What is the role of those regions? Parietal cortex support the 'first person perspective' on the visual world, unconsciously framing the visual object stream. Some prefrontal areas select and interpret conscious events for executive control. Such functions can be viewed as properties of the subject, rather than the object, of experience – the 'observing self' that appears to be needed to maintain the conscious state.

Humans seem to have a common intuition of an observing self that has access to conscious sensations, inner speech, images and thoughts. Philosophers such as Gilbert Ryle [1] denounced this idea as fallacious, but current evidence seems broadly supportive. This issue has become more pressing in the past decade as scientists have begun to revisit the basic topic of conscious experience. In brief, one can ask, does normal conscious experience involve an observing self?

Visual consciousness as a test case

Visual consciousness has been studied in depth and it is well established that visual features are identified in the ventral stream of posterior cortex. There, feature-sensitive cells support visual experiences of light, color, contrast, motion, retinal size, location and object identity; small lesions can selectively abolish those conscious properties [2]. However, to recall the experience of a human face, we need the hippocampal system¹. To respond to it emotionally, neurons in amygdala can be activated. But hippocampus and amygdala do not seem to support conscious contents directly. Thus, the ventral visual stream, which is

needed for specific conscious contents, seems to influence regions that are not (Box 1).

The direction of influence also goes the other way. When we step from a tossing sailboat onto solid ground, the horizon can be seen to wobble. On an airplane flight at night, passengers can see the cabin tilting on approach to landing, although they are receiving no optical cues about the direction of the plane. In those cases, unconscious vestibular signals (originating in the inner ear) shape conscious vision. In sum, conscious visual brain activities can influence unconscious ones, and vice versa.

Studying consciousness 'as such'

How do we know that conscious activity 'as such' evokes widespread regional interactions? After all, similar unconscious processes might do the same. Fortunately, a growing literature now compares the brain effects of conscious and unconscious stimulation. Precise experimental comparisons allow us to ask what conscious access does *per se*.

Many techniques permit comparisons between conscious and unconscious stimulation. In visual backward masking, a target picture is immediately followed by a scrambled image that does not block the optical input physically, but renders it unconscious. Binocular rivalry has been used for the same reason: it shows that when two competing optical streams enter the two eyes, only one consistent stream can be consciously perceived at any given moment. Most recently, several studies have demonstrated inattentional blindness, in which paying attention to one visual flow (e.g. a bouncing basketball) blocks conscious access to another activity at the very center of visual gaze (e.g. a man walking by in a gorilla suit). These studies generally show that unconscious stimuli still evoke local feature activity in sensory cortex² [3].

But what is the use of making something conscious if even unconscious stimuli are identified by the brain? Dehaene and colleagues have shown that although unconscious visual words activate known word-processing regions of visual cortex, the same stimuli, when conscious, trigger widespread additional activity in frontoparietal

Corresponding author: Bernard J. Baars (baars@nsi.edu).

¹ The hippocampal system evidently reflects conscious visual events with considerable fidelity. In a classic study of episodic memory (i.e. memory for conscious events), subjects were asked simply to pay attention to 10 000 distinct pictures over several days, with only five seconds of exposure per picture. On the following day they showed 96% recognition accuracy. No such memory feats are known for subliminal learning, suggesting that visual stimuli must be conscious in order to evoke spontaneous, highly efficient, episodic learning mediated by the hippocampal system [2,21].

² It is inherently difficult to prove the complete absence of consciousness in state studies. Sleep can vary in arousability from moment to moment, much like vegetative states and even general anesthesia. Some mentation is reported even from slow-wave sleep, and some waking-like functions can be preserved in rare brain damage patients who seem behaviorally unconscious. For most purposes, however, an absolute, stable zero point of consciousness is not needed. There is no question that deep sleep is much less conscious than full, responsive waking.

Box 1. Theoretical definitions

Conscious events can be defined in practice as those brain activities that subjects can report with high accuracy under optimal conditions, including minimal distraction and time delay. Unconscious events are those that are known to exist without the ability to report them accurately, such as subliminal activation of cortical color cells. Indeed, the word 'accurately reportable' could be used instead of 'conscious'. However, that would miss out something essential – namely, the fact that accurate reports are about experiences that all intact humans claim to have.

Global workspace theory is a cognitive architecture with an explicit role for consciousness [3,8,9]. Global workspace architectures have been studied in cognitive science, and have practical applications in organizing large, parallel collections of specialized processors, broadly comparable to the brain. In recent years, global workspace theory has been found increasingly useful by neuroscientists [3,4].

It makes minimal assumptions:

- That the brain can be viewed as a collection of distributed specialized networks, most of which do not directly support conscious experiences.
- That potentially conscious brain activities can compete for access to a neuronal global workspace capacity – a fleeting memory whose focal contents integrate multiple sources into a single, coherent brain representation, which is then widely distributed to many unconscious specialized networks. The transfer of information from conscious visual episodes to the (unconscious) hippocampal system

is a clear example of such distribution of conscious information in the brain.

- Prefrontal cortex has been suggested as one possible neuronal global workspace [4] but posterior sensory cortices also appear to integrate, briefly retain, and distribute coherent neuronal events that are reportable as conscious [9].
- That some unconscious networks, called contexts, are needed to shape conscious contents. Thus, contextual parietal maps of the visual field, which do not support conscious features, modulate visual feature cells that directly contribute to conscious aspects of seen objects.
- That such contexts can jointly constrain conscious events.
- That intentions and emotions can be viewed as 'goal contexts', shaping consciously reportable actions without themselves becoming conscious at the time.
- That hierarchies of goal contexts can act as executive networks to interpret and act upon conscious events without entering consciousness directly. Areas of prefrontal cortex appear to support such functions.

Franklin and colleagues have implemented global workspace theory in large-scale computer models, to make explicit predictions and test functionality in practical tasks (<http://csrg.cs.memphic.edu/>). Dehaene and colleagues have recently published a neural net model of dorsolateral prefrontal cortex in these terms [4].

regions [4,5]. This general result has now been replicated many times, using vision, touch, pain perception, and conscious versus automatic skills [3]. Together, these findings suggest that conscious access to a stimulus involves frontward spread of activation beyond the sensory regions of the posterior cerebrum.

Complementary findings come from studies of unconscious states. In deep sleep, auditory stimulation activates only primary auditory cortex [6]. In vegetative states following brain injury, stimuli that are ordinarily loud or painful activate only the primary sensory cortices [7,8]. Waking consciousness is apparently needed for forward spread of sensory activation to occur.

Global workspace theory and the brain

Global workspace theory emphasizes a two-way flow between conscious and unconscious brain activities [3,9–11]. The theory has been implemented in large-scale computational and neural net models [9–11] and bears a close resemblance to Neural Darwinist models [12]. However, it is helpful to think metaphorically of a theater of mind. In the conscious spotlight on stage – the global workspace – an actor speaks, and his words and gestures are distributed to many unconscious audience members, sitting in the darkened hall. Different listeners understand the performance in different ways. But as the audience claps or boos in response, the actor can change his words, or walk off to yield to the next performer. Finally, behind the scenes, an invisible (unconscious) director and playwright try to exercise executive control over the actor and the spotlight.

Less metaphorically, information appears to flow into a neuronal global workspace to be widely distributed. Such a structure must combine converging inputs – the actors competing for access to the spotlight – followed by momentary dominance of one coherent input, and then

wide distribution of output, in a wave of activity sent to other regions. In the brain, sensory projection areas could function much like a global workspace [3,9,10]. Some prefrontal regions play a role in selecting what enters consciousness (selective attention) and interpreting it to control voluntary action. This simple viewpoint helps to organize the evidence and generates testable hypotheses.

Context and the first-person perspective

Global workspace theory calls unconscious influences that shape conscious experiences 'contexts'. Parietal cortex does not recognize coffee cups; it has no feature cells for visual objects. But it does have unconscious egocentric (body-centered) and allocentric (object-centered) cellular maps, which shape our experience of coffee cups, paintings and our own bodies. Damage to right parietal cortex can cause contralateral neglect, a condition in which the left half of the visual field disappears from consciousness. Neglect patients cannot see the left side of a building they are looking at, and will only eat from the right half of a plate in front of their eyes. Thus the parietal region, which supports no reportable conscious activity in itself, can still profoundly shape conscious vision [13].

Neglect patients can also have disturbing alien experiences of their own bodies, especially of the left arm and leg. Such patients sometimes believe that their left leg belongs to someone else, often a relative, and can desperately try to throw it out of bed. Thus, parietal regions seem to shape contextually both the experience of the visual world and of one's own body. Notice that neglect patients still experience their alien limbs as conscious visual objects (a ventral stream function); they are just alien to oneself.

Vogeley and Fink [14] suggest that parietal cortex is involved in the first-person perspective, the viewpoint of the observing self. When subjects are asked to adopt the visual perspective of another person, functional

Table 1. Major properties of four types of unconscious state compared with conscious rest

State	Conscious resting state [19,20]	Deep sleep [22]	General anesthesia [24]	Vegetative state or coma [25]	Epileptic loss of consciousness [23]
Cause	Neuromodulation of the cortex by the brainstem, instructions to avoid deliberate tasks [26]	Physiological: neuromodulation of the forebrain by the brainstem	Pharmacological: a variety of chemical agents	Pathological: trauma, intoxication, anoxia, hypoglycemia	Pathological: slow, synchronized neuronal firing driven by brain foci [23]
Behavioral signs	Accurate reportability of attended stimuli; orientation to space, time, and self; visual images, inner speech, abstract thoughts; control of voluntary muscles	No reportability	No reportability	No reportability. Lower brainstem reflexes retained intact [25]	No reportability
Regional metabolism	High in frontoparietal cortex	Low in frontoparietal cortex [27]	Low in frontoparietal cortex [28]	Low in frontoparietal cortex	Low in frontoparietal cortex
EEG voltages	Low-amplitude, irregular high-frequency waves (8–100 Hz), or low-amplitude, regular alpha waves (8–12 Hz)	High-amplitude, regular, low-frequency waves (<4 Hz)	High-amplitude, regular, low-frequency waves	High-amplitude, regular, low-frequency waves	High-amplitude, spike-wave form, regular, low-frequency waves
Underlying neuronal mechanism (in cortical and thalamic neurons)	Firing irregularly at an average base rate of ~10 Hz	Slow, synchronized pausing of base-rate firing [22]	Slow, synchronous pausing of base-rate firing?	Slow, synchronous pausing of base-rate firing?	Slow, synchronous pausing of base-rate firing?
Functional connectivity	High and variable	Low between cortical regions, and between thalamus and cortex [22]	Low between cortical regions, and between thalamus and cortex	Low or absent between cortical regions, and between thalamus and cortex [29,30]	Low between cortical regions, and between thalamus and cortex [23]

Abbreviations: EEG, electroencephalogram.

magnetic-resonance imaging (fMRI) activity peaks in medial parietal, inferior lateral parietal and prefrontal cortex.

Prefrontal self systems

Parts of prefrontal cortex are believed to support other self functions. Damage there can change lifelong personality traits, such as the ability to inhibit antisocial impulses. The case of Phineas Gage is classical, and similar neurological patients are not uncommon [15]. Such personality functions seem contextual in that they rarely become conscious, and then only in passing. Yet they underlie the selection and interpretation of conscious thoughts, speech, emotion and social perception. They could constitute the point of view from which 'we' experience the world.

It is telling that patients with identity disorders such as fugue (a rapid change in personal identity lasting weeks or months) often show amnesia for the eclipsed self. When the patient returns to normal, he or she might report time loss – a period of weeks from which no conscious experiences can be recalled [16]. It is as if each personality serves to organize and interpret conscious events during its time of dominance. While parietal cortex appears to put visual scenes into context, prefrontal regions appear to do the same for more abstract aspects of experience, such as social, emotional and self evaluation.

Indeed, conscious experience in general can be viewed as information presented to prefrontal executive regions for interpretation, decision-making and voluntary control.

Crick and Koch note that 'it is useful to think of the front or higher/executive part of the cortex as looking at and interacting with the back, or sensory part.' [17] In split-brain patients, a great deal of executive control seems to require the 'narrative interpreter' of the speaking hemisphere. One reason is the pervasive role of inner and outer speech in daily life, to remind ourselves of things to do, to focus on current concerns, and to maintain verbally supported plans and beliefs.

These points suggest a new way to understand conscious and unconscious states.

Is the 'observing self' needed to sustain the conscious waking state?

Conscious waking shows fast oscillatory activity throughout the thalamocortical core, and functional connectivity that can change rapidly as a function of task, content and context [18] (Table 1). Despite this great neuronal variability, recent evidence suggests the existence of a conscious 'baseline state', a resting wakeful state in which no external tasks are required [19,20]. Surprisingly, when subjects are asked simply to rest, their frontoparietal metabolism is consistently higher than in standard cognitive tasks. This might be due to the flow of thoughts that subjects report in the resting state [20]. Spontaneous thoughts seem more self-relevant than standard cognitive tasks, which typically compel subjects to direct their attention away from their personal concerns.

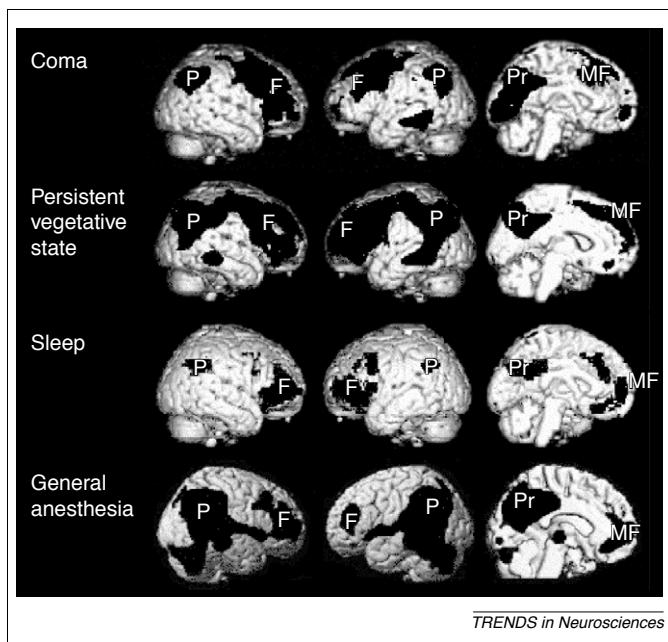


Figure 1. Metabolic activity in four types of unconsciousness, subtracted from conscious controls. Positron emission tomography (PET) scans showing regional decreases in metabolism or blood flow when unconscious states are compared with resting conscious states. Coma, persistent vegetative state, sleep and general anesthesia all show regional decreases in frontoparietal association cortices. Column 1 shows the right lateral aspect of the brain, column 2 the left lateral aspect, and column 3 a medial view of the left hemisphere. Abbreviations: F, prefrontal; MF, mesiofrontal; P, posterior parietal cortex; Pr, posterior cingulate/precuneus.

Metabolic activity in the conscious resting state is not uniformly distributed. Raichle *et al.* write that medial parietal regions, including ‘posterior cingulate cortex and adjacent precuneus can be posited as a tonically active region of the brain that may continuously gather information about the world around, and possibly within, us. It would appear to be a default activity of the brain...’ [19]. Mazoyer *et al.* also found high prefrontal metabolism during rest [20]. Notice that these are the same general areas that show additional activity when conscious sensory stimulation is compared with matched unconscious input. We will see that these regions also show markedly lower metabolism in unconscious states.

Four unconscious states: eclipsing the self?

Table 1 shows the conscious resting state compared with four unconscious states that are causally very different from each other: deep sleep³, coma/vegetative states, epileptic loss of consciousness⁴, and general anesthesia under six different anesthetic agents⁵.

³ At the level of cortical neurons, bursting rates do not change in deep sleep. Rather neurons pause together at <4 Hz between bursts [22]. Synchronous pausing could disrupt the cumulative high-frequency interactions needed for waking functions such as perceptual continuity, immediate memory, sentence planning, motor control and self-monitoring. It is conceivable that other unconscious states display similar neuronal mechanisms.

⁴ Although the spike-wave electroencephalogram (EEG) of epileptic seizures looks different from the delta waves of deep sleep and general anesthesia, it is also slow, synchronized and high in amplitude. The source and distribution of spike-wave activity varies in different seizure types. However, the more widespread the spike-wave pattern, the more consciousness is likely to be impaired [23]. This is again marked in frontoparietal regions.

⁵ Controversy exists over fast EEG oscillations under the influence of ketamine, and whether its effects should be viewed as primarily dissociative or anesthetic. Ketamine might be an exception to the slow-wave EEG found with other anesthetics.

Surprisingly, despite their very different mechanisms the four states share major common features. These include: (i) widely synchronized slow waveforms that take the place of the fast and flexible interactions needed for conscious functions⁵; (ii) frontoparietal regions becoming hypometabolic; (iii) widely blocked functional connectivity, both corticocortical and thalamocortical; and (iv) behavioral unconsciousness, including unresponsiveness to normally conscious stimuli. The first three of these features lower the probability of waking-type interactions among brain regions.

Figure 1 shows marked hypometabolism in the four unconscious states compared with conscious controls, precisely where we might expect: in frontoparietal regions. Could it be that brain regions that underlie the ‘observing self’ are thereby disabled?

Summary and future directions

Frontoparietal association areas have many functions beyond those touched on here. However, several lines of evidence suggest that they could have a special relationship with consciousness, even though they do not support the contents of sensory experience. (i) Conscious stimulation in the waking state leads to frontoparietal activation, but unconscious input does not; (ii) in unconscious states, sensory stimulation activates only sensory cortex, but not frontoparietal regions; (iii) the conscious resting state shows high frontoparietal metabolism compared with outward-directed cognitive tasks; and (iv) four causally very different unconscious states show marked metabolic decrements in the same areas.

Although alternative hypotheses must be considered, it seems reasonable to suggest that ‘self’ systems supported by these regions could be disabled in unconscious states. From the viewpoint of the observing self, this would be experienced as subjective loss of access to the conscious world. Unconscious states might not necessarily block the objects of consciousness; rather, the observing subject might not be at home.

Acknowledgements

B.J.B gratefully acknowledges support from the Neurosciences Institute and the Neurosciences Research Foundation (10640 John Jay Hopkins Drive, San Diego, CA 94549, USA; www.nsi.edu). S.L. is a research associate supported by the Belgian National Fund for Scientific Research (FNRS). We thank Björn Merker, Anil Seth, Douglas Nitz and E. Roy John for helpful discussions.

References

- 1 Ryle, G. (1949) *The Concept of Mind*, Hutchinson, London
- 2 Zeki, S. (2001) Localization and globalization in conscious vision. *Annu. Rev. Neurosci.* 24, 57–86
- 3 Baars, B.J. (2002) The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci.* 6, 47–52
- 4 Dehaene, S. *et al.* (2003) A neuronal network model linking subjective reports and objective physiological data during conscious perception. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8520–8525
- 5 Dehaene, S. *et al.* (2001) Cerebral mechanisms of word masking and unconscious repetition priming. *Nat. Neurosci.* 4, 752–758
- 6 Portas, C.M. *et al.* (2000) Auditory processing across the sleep–wake cycle: simultaneous EEG and fMRI monitoring in humans. *Neuron* 28, 991–999
- 7 Laureys, S. *et al.* (2000) Auditory processing in the vegetative state. *Brain* 123, 1589–1601

- 8 Laureys, S. *et al.* (2002) Cortical processing of noxious somato-sensory stimuli in the persistent vegetative state. *Neuroimage* 17, 732–741
- 9 Baars, B.J. (1988) *A Cognitive Theory of Consciousness*, Cambridge University Press
- 10 Franklin, S. (2000) Deliberation and voluntary action in ‘conscious’ software agents. *Neural Netw. World* 10, 505–521
- 11 Baars, B.J. and Franklin, S. (2003) How conscious experience and working memory interact. *Trends Cogn. Sci.* 7, 166–172
- 12 Edelman, G.M. (2003) Naturalizing consciousness: a theoretical framework. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5520–5534
- 13 Bisiach, E. and Geminiani, G. (1991) Anosognosia related to hemiplegia and hemianopia. In *Awareness of Deficit After Brain Injury: Clinical and Theoretical Issues* (Prigatano, G.P. and Schacter, D.L., eds), Oxford University Press
- 14 Vogeley, K. and Fink, G.R. (2003) Neural correlates of the first-person perspective. *Trends Cogn. Sci.* 7, 38–42
- 15 Anderson, S.W. *et al.* (1999) Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat. Neurosci.* 2, 1032–1037
- 16 American Psychiatric Association, (2000) *Diagnostic and Statistical Manual of Mental Disorders IV*, American Psychiatric Association
- 17 Crick, F. and Koch, C. (2003) A framework for consciousness. *Nat. Neurosci.* 6, 119–126
- 18 Buechel, C. and Friston, K. (2000) Assessing interactions among neuronal systems using functional neuroimaging. *Neural Netw.* 13, 871–882
- 19 Raichle, M.E. *et al.* (2001) A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682
- 20 Mazoyer, B. *et al.* (2001) Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Res. Bull.* 54, 287–298
- 21 Standing, L. (1973) Learning 10,000 pictures. *Q. J. Exp. Psychol.* 25, 207–222
- 22 Steriade, M. (2001) Active neocortical processes during quiescent sleep. *Arch. Ital. Biol.* 139, 37–51
- 23 Blumenfeldt, H. and Taylor, J. (2003) Why do seizures cause loss of consciousness? *Neuroscientist* 9, 1–10
- 24 John, E.R. *et al.* (2001) Invariant reversible QEEG effects of anesthetics. *Conscious. Cogn.* 10, 165–183
- 25 Laureys, S. *et al.* (1999) Cerebral metabolism during vegetative state and after recovery to consciousness. *J. Neurol. Neurosurg. Psychiatry* 67, 121
- 26 Hobson, J.A. and Pace-Schott, E.F. (2002) The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. *Nat. Rev. Neurosci.* 3, 679–693
- 27 Maquet, P. *et al.* (2000) Functional neuroimaging of normal human sleep by positron emission tomography. *J. Sleep Res.* 9, 207–231
- 28 Kaisti, K.K. *et al.* (2002) Effects of surgical levels of propofol and sevoflurane anesthesia on cerebral blood flow in healthy subjects studied with positron emission tomography. *Anesthesiology* 96, 1358–1370
- 29 Laureys *et al.* (2000) Restoration of thalamocortical connectivity after recovery from persistent vegetative state. *Lancet* 355, 1916
- 30 Laureys, S. *et al.* (1999) Impaired functional connectivity in vegetative state: preliminary investigation using PET. *Neuroimage* 9, 377–382

Trends Editorial Policy

Trends journals are indispensable reading for anyone interested in the life-sciences. At the heart of the journal there are authoritative Reviews and Opinions written in an accessible style by leading authors, which summarize exciting developments in your field. The majority of Reviews and Opinions are commissioned by the editor, but some originate from proposals. All of them are thoroughly peer reviewed.

- **Review** articles provide clear, concise, well illustrated and balanced views of recent advances, citing the primary literature and identifying important trends and key questions for on-going research.
- **Opinion** articles are special reviews designed to stimulate debate and cover controversial and emerging areas of research, and to present new models and hypotheses relating to a particular viewpoint.
- **Research Focus** articles provide detailed commentary and analysis in relation to recent primary research papers, and are restricted to 1500 words plus figures.



Global workspace dynamics: cortical “binding and propagation” enables conscious contents

Bernard J. Baars^{1*}, Stan Franklin² and Thomas Zoega Ramsøy³

¹ The Neurosciences Institute, La Jolla, CA, USA

² University of Memphis, Memphis, TN, USA

³ Copenhagen University, Copenhagen, Denmark

Edited by:

Melanie Boly, Belgian National Fund for Scientific Research, USA

Reviewed by:

David Rosenthal, City University of New York, USA

Guido Hesselmann, Charité Campus Mitte, Germany

Murray Shanahan, Imperial College London, UK

***Correspondence:**

Bernard J. Baars, The Neurosciences Institute, 800 Silverado Street, Suite 302, La Jolla, CA 92037, USA.
e-mail: baarsbj@gmail.com

A global workspace (GW) is a functional hub of binding and propagation in a population of loosely coupled signaling elements. In computational applications, GW architectures recruit many distributed, specialized agents to cooperate in resolving focal ambiguities. In the brain, conscious experiences may reflect a GW function. For animals, the natural world is full of unpredictable dangers and opportunities, suggesting a general adaptive pressure for brains to resolve focal ambiguities quickly and accurately. GW theory aims to understand the differences between conscious and unconscious brain events. In humans and related species the cortico-thalamic (C-T) core is believed to underlie conscious aspects of perception, thinking, learning, feelings of knowing (FOK), felt emotions, visual imagery, working memory, and executive control. Alternative theoretical perspectives are also discussed. The C-T core has many anatomical hubs, but conscious percepts are unitary and internally consistent at any given moment. Over time, conscious contents constitute a very large, open set. This suggests that a brain-based GW capacity cannot be localized in a single anatomical hub. Rather, it should be sought in a *functional hub* – a dynamic capacity for binding and propagation of neural signals over multiple task-related networks, a kind of neuronal cloud computing. In this view, conscious contents can arise in *any* region of the C-T core when multiple input streams settle on a winner-take-all equilibrium. The resulting conscious gestalt may ignite an any-to-many broadcast, lasting ~100–200 ms, and trigger widespread adaptation in previously established networks. To account for the great range of conscious contents over time, the theory suggests an open repertoire of binding¹ coalitions that can broadcast via theta/gamma or alpha/gamma phase coupling, like radio channels competing for a narrow frequency band. Conscious moments are thought to hold only 1–4 unrelated items; this small focal capacity may be the biological price to pay for global access. Visuotopic maps in cortex specialize in features like color, retinal size, motion, object identity, and egocentric/allocentric framing, so that a binding coalition for the sight of a rolling billiard ball in nearby space may resonate among activity maps of LGN, V1-V4, MT, IT, as well as the dorsal stream. Spatiotopic activity maps can bind into coherent gestalts using adaptive resonance (reentry). Single neurons can join a dominant coalition by phase tuning to regional oscillations in the 4–12 Hz range. Sensory percepts may bind and broadcast from posterior cortex, while non-sensory FOKs may involve prefrontal and frontotemporal areas. The anatomy and physiology of the hippocampal complex suggest a GW architecture as well. In the intact brain the hippocampal complex may support conscious event organization as well as episodic memory storage.

Keywords: consciousness, awareness, attention, voluntary control, global workspace theory, cortex, theoretical biology, brain dynamics

INTRODUCTION

This paper extends the Global Workspace (GW) theory of conscious experience to brain evidence, particularly the role of the

cortex and thalamus (Baars, 1988, 2002; Dehaene and Naccache, 2001; Baars and Franklin, 2003, 2007; Dehaene et al., 2006; Shanahan, 2010, 2012; Dehaene and Changeux, 2011; etc.). While cortex

¹The term “binding” was originally applied to the integration of many neural sources of sensory features like color, visual field location, size, etc., and at higher levels object and face identity, semantic category, and the like. As evidence for higher level binding has built up, it is now also applied to the integration of

brief events, perceived causality, and self-object integration. All conscious contents are internally consistent, as pointed out here, and combine numerous receptor inputs into single, coherent, and voluntarily reportable objects or events.

and thalamus look separate to the naked eye, they act as an integrated system (Llinas and Pare, 1991; Edelman and Tononi, 2000; Steriade, 2006; Freeman, 2007). A GW approach helps to organize a large body of evidence and yields distinctive predictions².

High quality empirical evidence has increased rapidly in recent years, both for conscious state studies and for conscious contents. State studies typically compare waking to slow-wave sleep, coma, general anesthesia, and the epilepsies. Studies of conscious contents compare conscious vs. unconscious cognition during the waking state, using binocular rivalry, the attentional blink, backward masking, and attentional manipulations. Both conscious and unconscious stimuli trigger sensory volleys that can be traced well into the cortex (Gaillard et al., 2009; Panagiotaropoulos et al., 2012).

Global Workspace theory has been applied to numerous empirical findings (e.g., Baars, 1988, 2002; Madl et al., 2011, 2012). Brain imaging experiments have supported the best-known GW prediction of “widespread integration and broadcasting” (Dehaene and Naccache, 2001). That is, conscious stimuli typically evoke cortical activity that is more widespread, intense, and correlated than matched unconscious stimuli.

Some of the best evidence to date comes from Panagiotaropoulos et al. (2012) building on almost two decades of findings from intracranial recordings in the macaque. They used an experimental technique called “flash suppression,” involving a long-lasting type of binocular rivalry between the conscious (perceived) but not sensory (unperceived) visual channel. This method allows for “contrastive analysis” of conscious vs. unconscious contents with identical stimulus presentation to the two eyes. Previous work from the Logothetis laboratory has shown conscious visual stimulus representation in temporal regions including IT, MTL, and the temporal association cortex. The most recent report extends this to the lateral prefrontal cortex reporting, “neuronal discharges in the majority of single units and recording sites in the LPFC follow the phenomenal perception of a preferred stimulus. Furthermore, visual awareness is reliably reflected in the power modulation of high-frequency (>50 Hz) local field potentials in sites where spiking activity is found to be perceptually modulated.”

While conscious perception appears to be markedly more distributed and in higher signal fidelity as compared to matched unconscious contents, this does not necessarily apply to other varieties of cognition. Unconscious automatic skills and perceptual inferences, parietal ego- and allocentric maps, sensorimotor control and implicit memory retrieval may show accuracy equal to, or greater than matched conscious events. The brain basis for these differences is not well understood.

This paper develops GW dynamics, suggesting that conscious experiences reflect a flexible “binding and broadcasting” function in the brain, which is able to mobilize a large, distributed collection of specialized cortical networks and processes that are not conscious by themselves. Note that the “broadcast” phase proposed by the theory should evoke widespread adaptation, for the

same reason that a fire alarm should evoke widespread responding, because the specific needs for task-relevant responders cannot be completely known ahead of time. General alarms are interpreted according to local conditions.

A brain-based GW interacts with an “audience” of highly distributed, specialized knowledge sources, which interpret the global signal in terms of local knowledge (Baars, 1988). The global signal triggers reentrant signaling from receiving networks, allowing for increasingly focused resonance between source networks and receiving networks.

DYNAMIC GW VIS-À-VIS OTHER THEORETICAL PROPOSALS

We can widely divide current theories into philosophical and empirically based ones. Some of the philosophical theories are now generating testable hypotheses.

Empirical theories can be divided into “localist” vs. “local-global” types. There are no exclusively global theories, since no one denies the evidence for local and regional specialization in the brain.

Philosophical theories typically aim to account for subjective experiences or “qualia,” a notoriously difficult question. Recently some philosophical perspectives, like “higher order theory” (HOT) have also generated testable proposals about the involvement of brain regions like the prefrontal cortex. However, brain imaging experiments (e.g., Dehaene and Naccache, 2001) have long implicated the frontoparietal cortex in subjective experience.

It is not clear at this time whether philosophically based theories generate novel, testable predictions. However, efforts are under way to test HOT theories. In general, claims to explain subjective qualia are still debated.

Empirical theories tend to start with reliable phenomena (See Appendix A). For example, the limited capacity of conscious perceptual contents has been discussed since the nineteenth century, and continues to amass a large basis in reliable evidence. One of the driving questions of GW theory is how the limited capacity of momentary conscious contents can be reconciled with the widespread access enabled by conscious contents.

Zeki (2001) makes the localist claim that conscious percepts of red objects involve “micro-conscious” activation of cortical color regions (visual areas V3/V4). However, most empirical theories combine local and global activities, as briefly discussed above. It is still possible that momentary events may be localized for 100 ms or less, and that full conscious contents emerge over some hundreds of milliseconds.

The Dynamic GW theory (dGW) is a specific version of the “dynamic core” hypothesis proposed by Edelman and Tononi (2000), and in somewhat different forms, by Edelman (1989) and others. The cognitive basis of the theory was elaborated by Baars (1988) and in subsequent papers with Franklin and others. Additionally, Dehaene, Changeux, and many coworkers have worked out specific, testable experimental models.

Dynamic Global Workspace theory implies a directional signal flow from binding to receiving coalitions. For each conscious event there is a dominant source and a set of receivers, where the propagated signal is interpreted, used to update local processes, and refreshed via reentrant signaling to the source (Edelman, 1989). Conscious sensations arise in a different center of binding and

²Previous work is available at Cognitive Computing Research Group, The University of Memphis (<http://ccrg.cs.memphis.edu/papers.html>) Steriade (2006) we use “C-T” to emphasize the driving role of the cortex.

propagation than “feelings of knowing (FOK),” like the tip-of-the-tongue (TOT) experience, as demonstrated by brain imaging studies (Maril et al., 2001). Directional broadcasting of bound conscious contents is one testable distinction from other proposals (Edelman et al., 2011). Supportive evidence has been reported by Doesburg et al. (2009) and others.

Other theories, like Tononi’s mathematical measure of complexity, phi, seem less directional (Edelman and Tononi, 2000). Llinas and Pare (1991) have emphasized the integration of specific and non-specific thalamocortical signaling, and Freeman et al. (2003) have developed a conception of hemisphere-wide signaling and phase changes.

Nevertheless, current local-global theories are strikingly similar. Whether major differences will emerge over time is unclear.

DYNAMIC GW AS A LOCAL-GLOBAL THEORY

In 1988 GW theory suggested that “global broadcasting” might be one property of conscious events. Other proposed properties were:

1. Informativeness, that is, widespread adaptation to the novelty of the reportable signal, leading to general habituation (information reduction) of the news contained in the global broadcast. The evidence now supports widespread neuronal updating to novel input.
2. Internal consistency of conscious contents, because mutually exclusive global broadcasts tend to degrade each other. This is a well-established feature of conscious contents, first observed in the nineteenth century and replicated many thousands of times. Binocular rivalry is one well-known example.
3. Interaction with an implicit self-system. Baars (1988) proposed that the observing self is coextensive with implicit frames that shape the objects of consciousness. The notion of an implicit self now has a great deal of psychological evidence in its favor. Lou et al. (2010) have shown that it may involve the precuneus and orbitofrontal cortex. Others have argued for a midline brain system ranging from the PAG to the orbitofrontal cortex.
4. Limited capacity, and an operating cycle of 100–200 ms. The functional reason for conscious limited capacity has not yet received a satisfactory explanation. We will sketch a possible account below.

After almost a century of scientific neglect, many other aspects of consciousness remain to be studied. In recent times two research programs have pursued GW theory (GWT) in detail. A French group led by Dehaene and Changeux have conducted almost two decades of important research, focusing on formal models of specific experimental paradigms that allow close comparisons between conscious and unconscious conditions. This has been the most systematic and focused modeling and testing program on GWT, called “neuronal GW theory” by the authors. It continues to increase in sophistication and empirical reach over time (e.g., Dehaene and Naccache, 2001).

Franklin and coworkers have pursued GWT in terms of a large-scale cognitive architecture by (Baars and Franklin, 2003, 2007; Snaider et al., 2011; Franklin et al., 2012). A series of published papers by this group suggest a role for conscious events in working

memory, selective attention, executive functions, FOK, problem-solving, the memory systems and broader biological functions (see below).

Because almost all neural links in the cortico-thalamic (C-T) system are bidirectional, reentrant signaling from receivers to broadcasting sources may quickly establish task-specific signaling pathways, in the same way that a fire department might locate the source of a community-wide alarm, and, then, communicate in a much more task-specific way. Current evidence suggest brief broadcasts, as suggested by the ~100 ms conscious integration time of different sensory inputs.

THE REPORTABILITY OF CONSCIOUS EVENTS

Scientific theories often begin with intuitively plausible measures, but in time they must explain those measures. “Accurate report” has long been the most widely used behavioral index of conscious experiences (Baars, 1988). All of sensory psychophysics is based on accurate report, but we have no accepted explanation why conscious cognition should be reportable, while unconscious cognition is not. We suggest an answer to this puzzle below.

CORTEX AND THALAMUS

Many brain regions have been proposed to underlie conscious experiences, including the cortex, thalamus, brainstem reticular formation, claustrum, zona incerta, colliculi, prefrontal cortex, visual feature fields, thalamocortical projections, and the like. Recent evidence (e.g., Damasio, 1989; Edelman, 1989; Steriade, 2006; Freeman, 2007; and others) from recordings in the living brain support the “cortex plus thalamus” account for conscious contents in the waking state, at least in humans and other primates (Figure 1). Non-primates may utilize analogous brain structures, like the avian pallium.

This C-T view is based on:

- (a) Behavioral and brain imaging studies of cognition during waking compared to sleep, drowsiness, distraction, extended mental effort, sleep deprivation, coma, general anesthesia, and other states.
- (b) Direct intracranial stimulation and recording during specific cognitive tasks, like speaking vs. listening to a word.
- (c) Empirical dissociations between conscious vs. unconscious brain conditions, like split brain surgery, cortical blindness and parietal neglect. These phenomena can often be replicated in healthy subjects using TMS, hypnosis, drugs, or behavioral manipulations.

The case for the C-T system has recently been strengthened by the use of TMS-EEG, a magnetic pulse aimed at a specific cortical location while recording from a neurally linked one (Massimini et al., 2005). For example, TMS applied to the left somatosensory hand region leads to strong EEG pickup in the contralateral hand region, much like a conductance meter in an electrical circuit. TMS-EEG may become the first practical, minimally invasive measure of causal connectivity in cortex. It has been shown to differentiate between conscious waking vs. sleep and coma, and to distinguish behavioral coma from true coma (Rosanova et al., 2012).

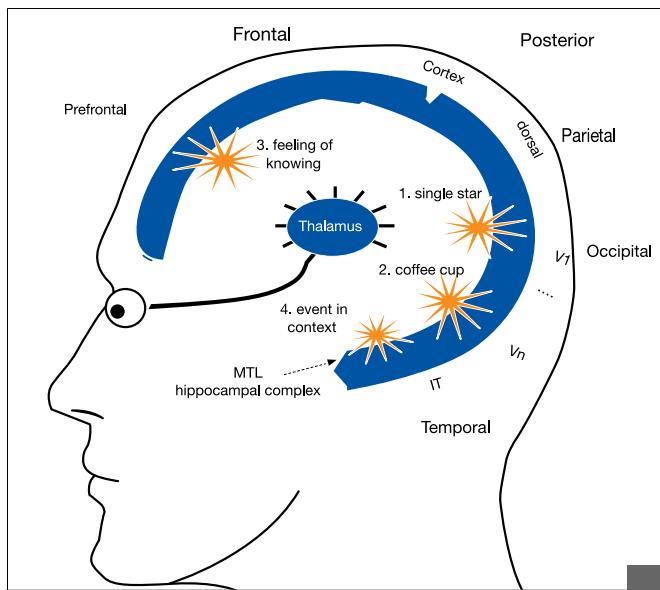


FIGURE 1 | Binding and broadcasting from many locations. Four alternative sources of binding and broadcasting in the C-T core. Sites of possible binding and broadcasting are shown as yellow starbursts. Because global broadcasts mutually interfere, only one can occur in any 100–200 ms period. Global interference may explain the limited capacity of momentary conscious contents. Areas V1 and IT: visuotopic maps include area V1, the first cortical map for vision, and area IT, where conscious object representations emerge (Logothetis, 1998). Different coalitions of anatomically identical spatiotopic activity maps may lead to different conscious experiences, like the sight of a single star on a dark night (V1) vs. a coffee cup within arm's reach (IT). Prefrontal cortex: non-sensory "feelings of knowing" may bind and broadcast from non-sensory cortex. Area MTL: we predict that the intact medial temporal cortex contributes to subjective event organization as well as episodic memory coding (See The Hippocampus and Conscious Contents: a Novel Prediction).

Damage to frontoparietal regions or connections may impair the conscious state. Demertzi et al. (2012) and Blumenfeld (2012) have therefore proposed a consciousness system including large regions of frontal and parietal cortex. While this proposal helps to explain waking state impairments, it does not include the classical sensory cortices. However, sensory events are often thought to be the prototypical contents of consciousness.

We suggest below that in the intact brain, sensory regions can drive frontoparietal areas, so that both regions play essential roles. More generally we propose that conscious events may arise anywhere in the C-T complex during waking, and that they may signal many other loci in the C-T core (see Figure 1). This proposed "any-to-many" recruitment is testable with current methods.

In many cases the original binding source specifies the qualitative conscious contents (the qualia), so that visually attributed experiences are expected to be dominantly sourced in visual cortex, while simple sounds arise in auditory regions. As we will see, non-sensory regions of the cortex can give rise to FOKs, which are richly endowed with reportable contents, even in the absence of sensory cues.

In this view conscious experiences reflect flexible "binding and broadcasting" in the C-T complex, interacting with subcortical,

striate and cerebellar regions. GW dynamics (dGW) emphasizes the range and flexibility of binding and broadcasting. It may also help to explain the feats of functional adaptation that can occur after severe cortical damage (see Voluntary Reports of Conscious Events).

Here we aim to preserve the simplicity of the theory while taking into account a wider range of evidence.

Cortex and thalamus consist of multi-layered two-dimensional arrays of neurons and their connective arborizations. Each array projects topographically to others. Under the cortical mantle, white matter pathways run in all canonical directions. These long pathways are targeted in very precise and regular topographies. Local cortical links are also very numerous. Tractographic studies indicate that the connectivity of the C-T system follows small-world network properties (Sporns and Honey, 2006).

The C-T core constitutes by far the biggest parallel-interactive structure in the mammalian brain, showing massive wave-like activities during active states (Steriade, 2006). Most sensory pathways flow into the C-T system, while motor pathways flow outward for subcortical, craniospinal, and autonomic control.

Hundreds of specialized regions of cortex have been identified, and new ones continue to be discovered. Cortical regions are mirrored in major thalamic nuclei, most of which are also layered arrays. Sherman and Guillery (2011) note that "any new information reaching a cortical area... benefits from a thalamic region." Thalamic nuclei do not communicate with each other directly, but only by way of cortical synapses.

Cortico-thalamic arrays are mutually linked via spatially consistent "labeled lines." For example, the small foveal center of each retina subtends only 2–4° of visual arc, with about one million densely packed cones and rods. Foveal patches can be roughly considered to be 1000×1000 arrays of light receptors that are echoed point-to-point in retinal ganglion cells, whose axons make up the optic nerve. Ganglion cells are mirrored in the visual thalamus (LGN), which transmit signals point-to-point to V1.

Labeled line coding implies that the relative locations of neurons are preserved in higher level retinotopic maps. Beyond V1 visuotopic maps preserve a sparser copy of the visual field, while cellular receptive fields increase in size and decrease in spatial resolution. Spatiotopic mapping is regular and systematic throughout the C-T complex. For instance, there are tonotopic maps in auditory areas, body maps in the somatic regions, and similar such maps in the motor cortex. At higher levels C-T arrays project to self-centered and object-centered spaces, which must be coordinated with each other to support a coherent domain of action and perception. Multiple levels of inhibition serve to regulate and sculpt the excitatory activity of the cortex and its satellites. Cortical arrays therefore resemble the head-up cockpit display of a fighter jet, which may show a geometric frame for the plane itself, another frame for targets, and more for attackers and neutrals. When these floating spatial frames intersect, they may alert executive networks to make decisions.

Starting with LGN, all spatiotopic pathways become bidirectional. Successive arrays pick up visual features like spatial frequency, contrast, edge orientation, gestalt properties, hue, motion, and object identity. Higher level properties like object permanence, size constancy, color constancy, shape from shading, face

and object recognition, scene analysis, movements, causality, and event organization, all require complex interactions among 40 or more spatiotopic arrays. The sight of a red traffic light must remain stable in spite of differences in reflectance, observer motion, background clutter, and changes in sunlight.

Figure 1 shows four examples of possible binding and broadcasting in the C-T core (starburst icons). Cortical area V1 and the visual thalamus (LGN) can be conceived as two arrays of high-resolution bright and dark pixels, without color. The sight of a single star on a dark night may therefore rely heavily on V1 and its mirror array of neurons in LGN. V1 and LGN interact constantly, with bidirectional signal traffic during waking.

The conscious sight of a single star at night reveals some surprising features of conscious vision, including spatial context sensitivity, as in the classical autokinetic effect: single points of light in a dark space begin to wander long subjective distances in the absence of spatial framing cues. The autokinetic effect is not an anomaly, but rather a prototype of decontextualized percepts (Baars, 1988). A large literature in perception and language shows scores of similar phenomena, as one can demonstrate by looking at a corner of a rectangular room through a reduction tube that excludes external cues. Any two or three-way corner in a carpentered space is visually reversible, much like the Necker Cube and the Ames trapezoid. Such local ambiguities exist at every level of language comprehension and production (Baars, 1988; Shanahan and Baars, 2005).

The dorsal stream of the visual cortex provides egocentric and allocentric “frames” to interpret visual events in nearby space. These parietal frames are not conscious in themselves, but they are required for visual objects to be experienced at all (Goodale and Milner, 1992). Injury to right parietal cortex may cause the left half of visual space to disappear, while contralateral stimulation, like cold water in the left ear, may cause the lost half of the field to reappear.

Thus even a single dot of light in a dark room reveals the contextual properties of conscious perception. Ambiguity and its resolution is a universal need for sensory systems in the natural world, where ambiguity is commonly exacerbated by camouflage, deceptive signaling, distraction, unpredictable movements, ambushes, sudden dangers and opportunities, darkness, fog, light glare, dense obstacles, and constant utilization of cover by predators and prey (Bizley et al., 2012).

Conscious percepts plausibly involve multiple “overlays,” like map transparencies. The sight of a coffee cup may involve an object overlaid by color, texture, and reflectance, combining information from LGN, V1, V2, V3/V4, and IT (Crick and Koch, 1990). Active cells in those arrays may stream signals across multiple arrays, cooperating, and competing to yield a winner-take-all coalition. Once the winning coalition stabilizes, it may “ignite” a broadcast to other regions.

Conscious vision is strikingly flexible with respect to level of analysis, adapting seamlessly from the sight of a single colored dot to the perception of a dotted (pointillist) painting. An account of conscious vision must therefore explain how a local dot can be perceived in the same visual display as a Georges Seurat painting³.

To identify a single star at night, because the highest spatial resolution is attained in the retina, LGN, and V1, the visual cortex must be able to amplify neuronal activity originating in LGN-V1 through attentional modulation. For coffee cups and faces, the relative activity of IT and the fusiform gyrus must be increased. It follows that binding coalitions of visual activity maps can bring out the relative contribution of different feature levels, even for the same physical stimulus (Itti and Koch, 2001).

BIDIRECTIONAL PATHWAYS AND ADAPTIVE RESONANCE

Because C-T pathways are bidirectional they can support “reentrant signaling” among topographically regular spatial maps. The word “resonance” is often used to describe C-T signaling (Wang, 2001). It is somewhat more accurate than “oscillation,” which applies to true iterative patterns like sine waves. Edelman and coworkers prefer the term “reentry,” while others, like use “adaptive resonance.” We will use the last term to emphasize its flexible, selective, and adaptive qualities.

Adaptive resonance has many useful properties, as shown in modeling studies like the Darwin autonomous robot series, where it can account for binding among visual feature maps, a basic property of visual perception (**Figure 2**) (Izhikevich and Edelman, 2008). Edelman has emphasized that reentry (adaptive resonance) is not feedback, but rather evolves a selectionist trajectory that can

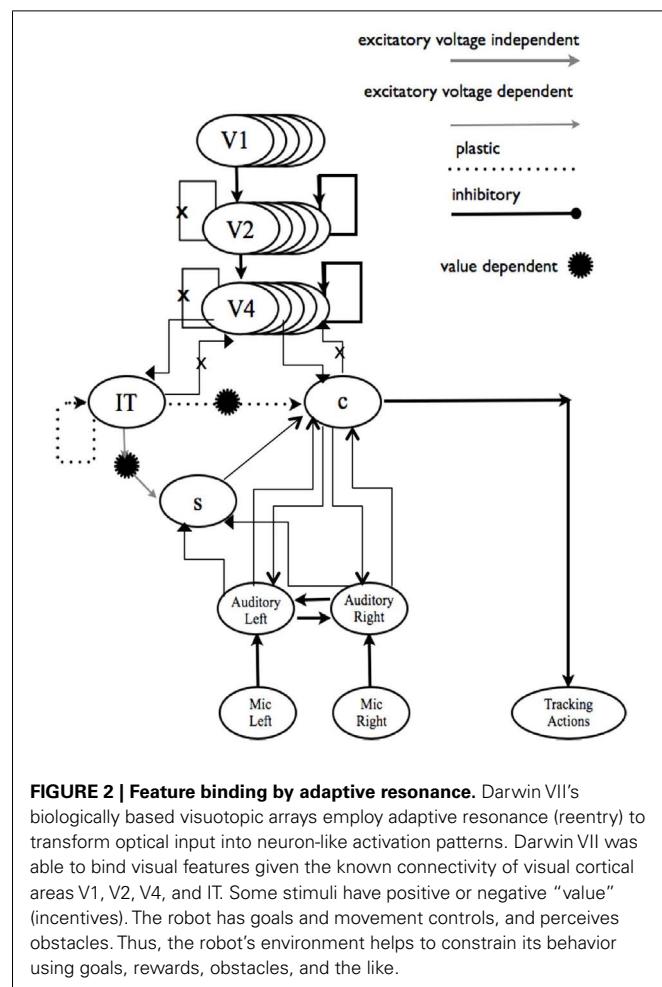


FIGURE 2 | Feature binding by adaptive resonance. Darwin VII's biologically based visuotopic arrays employ adaptive resonance (reentry) to transform optical input into neuron-like activation patterns. Darwin VII was able to bind visual features given the known connectivity of visual cortical areas V1, V2, V4, and IT. Some stimuli have positive or negative “value” (incentives). The robot has goals and movement controls, and perceives obstacles. Thus, the robot's environment helps to constrain its behavior using goals, rewards, obstacles, and the like.

³See <http://en.wikipedia.org/wiki/Pointillism>

search for solutions to biologically plausible problems. Grossberg and others have developed adaptive resonance models for cortical minicolumns and layers.

A variety of high-level wave-like phenomena emerge in the C-T system, including standing and traveling waves, spiral vortices, centrifugal propagation, phase coupling and decoupling, microstates, cross-frequency coupling, and hemisphere-wide phase transitions at theta-alpha rates (Freeman et al., 2003; Izhikevich and Edelman, 2008). Complex waveforms in the core range from 0.1 to 200 Hz, with momentary spikes up to 600 Hz.

The basic unit of the C-T is therefore not the single neuron, nor the traditional one-directional sensory pathway. It is rather “a unit of adaptive resonance,” which can be thought of as an artificial neural network with at least two layers. Bidirectionality makes the C-T core different from the cerebellum and basal ganglia, which do not support conscious contents directly. Damage to the cerebellum does not directly impair conscious contents, though it can devastate fine motor control. The cerebellum has parallel modules and pathways comparable to a computer server farm, while the C-T complex is parallel-interactive, so that any array of neurons can signal any other. While the cerebellum is more “cognitive” than previously thought, it does not directly enable conscious contents.

Most C-T activity is endogenous: “the cortex talks mostly to itself.” With a few exceptions thalamic nuclei do not communicate with each other directly, but are driven by cortical regions. Steriade (2006) concluded, “*The cerebral cortex and thalamus constitute a unified oscillatory machine displaying different spontaneous rhythms that are dependent on the behavioral state of vigilance.*” (Italics added.)

This is not to minimize the role of neuronal spike timing, which is known to evoke synaptic plasticity. The C-T core may use multiple neuronal codes, including non-classical propagation via glial cells, electromagnetic field induction, electrical synapses, and even membrane ion currents. In conventional terms, brain rhythms interact with single cell activity. Population rhythms require interacting single cells that are both excitatory and inhibitory. Regional rhythms like theta can also recruit single cells that phase-adapt to the peak of the regional wave (Canolty et al., 2010).

Thus neural coding is simultaneously spatial and temporal. The C-T nexus appears to be a vast, parallel-interactive signaling medium, capable of local, regional, and global processing. Tononi has proposed that only the C-T complex combines high integrative capacity within perceived events along with high differentiation between them.

Bidirectional signaling between linked arrays supports many kinds of emergent signaling, just as random Brownian motion allows for many kinds of acoustical wave propagation. Izhikevich and Edelman (2008) have shown that an accurate neuronal simulation of the cortex gives rise to a range of macroscopic waveforms from <0.1 to 200 Hz. Many oscillatory phenomena have been observed in the cortex, including traveling waves, 4–7 Hz microstates, vortices, and centrifugal propagation (Freeman, 2009).

FEATURE AND FRAME BINDING

Visual features are stimulus properties that we can point to and name, like “red,” “bounded,” “coffee cup,” “shiny,” etc. Feature binding is a well-established property of sensory perception.

There is much less discussion about what we will call “frame binding,” which is equally necessary, where “frames” are defined as visual arrays that do not give rise to conscious experiences, but which are needed to specify spatial knowledge within which visual objects and events become conscious. Powerful illusions like the Necker Cube, the Ames trapezoidal room, the railroad (Ponzo) illusion are shaped by unconscious Euclidian assumptions about the layout of rooms, boxes, houses, and roads.

Unconscious assumptions that shape conscious features were previously called “contexts” (Baars, 1988). Because that term is now used with other meanings in the psychological literature, we will adopt to word “framing,” which is in common use in the social sciences. The best-known brain examples are the egocentric and allocentric visuotopic arrays of the parietal cortex. When damaged on the right side, these unconscious visuotopic fields cause the left half of objects and scenes to disappear, a condition called hemi-neglect. Goodale and Milner have shown that even normal visuomotor guidance in near-body space may be unconscious. In vision the dorsal “framing” stream and “feature-based” ventral stream may combine in the medial temporal cortex (MTL) (Shimamura, 2010, see The Hippocampus and Conscious Contents: A Novel Prediction below). Baars (1988) reviewed extensive evidence showing that unconscious framing is needed for normal perception, language comprehension and action planning. In sum, normal conscious experiences need both traditional feature binding and frame binding (Shanahan and Baars, 2005)⁴.

CONSCIOUSNESS ENABLES MANY KINDS OF ACCESS (FIGURE 3)

Animals live in a world of unknowns, surrounded by dangers, and opportunities that may be fleeting, hidden, camouflaged, surprising, deceptive, and ambiguous. Conscious brains may have evolved to cope with such unknowns (Baars, 1988, 2002; Baars, 1988). Newell and colleagues built the first GW architecture to perform acoustical word recognition, at a time when that task was largely underdetermined (Newell, 1990). Their solution was to build a computational architecture, a blackboard model, which would allow many incomplete sources to compete and cooperate to resolve some focal ambiguity. The result was remarkably successful for its time in recognizing nearly 1,000 ordinary words spoken in normal acoustical spaces, complete with hard echoing surfaces, mumbling speakers, and soft, absorbent surfaces, background noises, and the like. Speech recognition is now handled with improved formant tracking, but even today, if semantic unknowns arise in a spoken word stream, a GW architecture may be useful to find the answer. We have no semantic algorithms that interpret word ambiguities across many domains, the way humans routinely do.

Baars and Franklin (2003) used GW theory to propose that consciousness enables access between otherwise separate knowledge sources. One major kind of access that has been discussed since Emmanuel Kant is the access of the “observing self” to the contents of consciousness. Lou et al. (2010) have shown that self-related brain regions like the precuneus and midline structures from

⁴To avoid confusion, this paper uses the word “frame” instead of “context” for unconscious constraints on conscious contents (Baars, 1988).

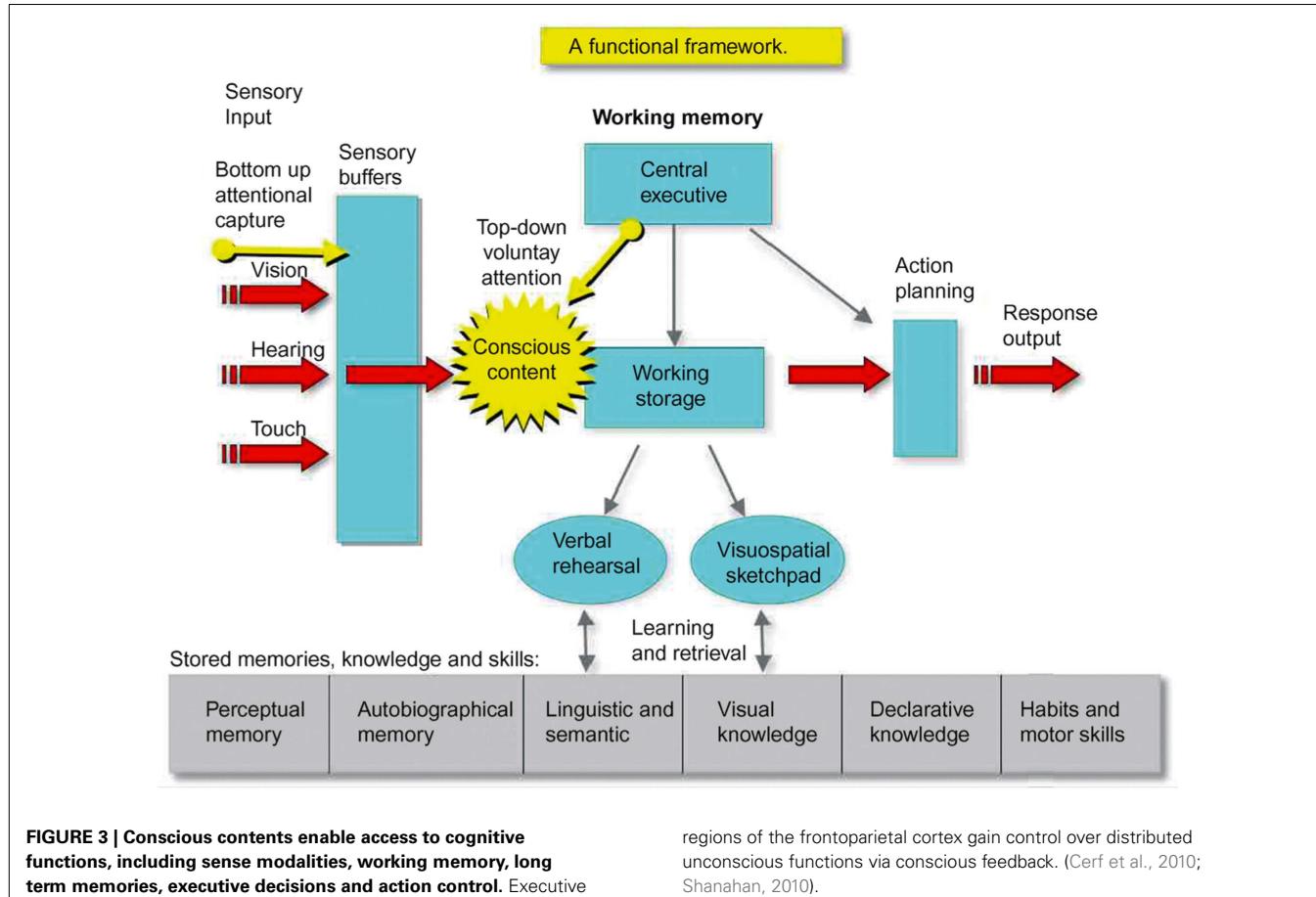


FIGURE 3 | Conscious contents enable access to cognitive functions, including sense modalities, working memory, long term memories, executive decisions and action control. Executive

regions of the frontoparietal cortex gain control over distributed unconscious functions via conscious feedback. (Cerf et al., 2010; Shanahan, 2010).

the PAG to orbitofrontal cortex may be mobilized by conscious sensory contents. Baars (1988) proposed that self-other access is a specific variety of framing (contextualizing), and that it is a necessary condition for conscious contents.

Global workspace architectures can also “call” fixed automatisms. For example, in speech recognition word ambiguity may be resolved by a known syntactic rule. A global broadcast of the ambiguous word may recruit routines whose relevance cannot be known ahead of time. We have referred to this as contextualization or frame binding (Baars, 1988; Shanahan and Baars, 2005⁵⁶⁷). The “frame problem” is a recognized challenge in artificial intelligence and robotics, but it applies equally to living brains.

Animals and plants co-evolve both accurate and deceptive signals, as in the case of monkeys, bees and flowers. In such an evolutionary context the need to deal with ever-changing unknowns may result in a constant adaptive pressure toward certain kinds of brains. The C-T system occupies 80% of the cranial volume in humans, mostly connective pathways between thin layers of neuronal cell bodies. This costly organ must somehow improve adaptive fitness (Baars, 2002). Specialized algorithms like spinal reflexes are crucial, but in the face of constantly evolving unknowns, a GW architecture may provide a competitive edge. Here, we consider how such a flexible adaptive capacity may be expressed in the brain.

SPATIOTOPIC ACTIVITY MAPS, STREAMS, AND COALITIONS

Spatiotopic organization is common in the C-T core, including the tonotopic maps of the auditory system, multimodal maps of parietal cortex, and visual maps of the prefrontal cortex. Both sensory

⁵The LIDA model suggests that cognitive processes involve cascading action-perception cycles (Franklin et al., 2012), called cognitive cycles in the model. After filtering for salience, each cognitive cycle produces a global broadcast that both facilitates several modes of learning and recruits resources for consciously mediated action selection. During the preconscious understanding phase of the cognitive cycle, the current situational model (Franklin et al., 2012), acting as a blackboard model, is updated by accessing previously learned information from sensory memory, recognition memory, spatial memory, episodic memory and semantic memory. The next contents of consciousness selects itself dynamically from cognitive structures whose building blocks are the accessed cognitive representations. Thus consciousness creates access leading to conscious thoughts and subsequent actions.

⁶The term “oscillation” is often used to describe wave-like activity in the brain. However, Freeman (2007) points out that there is no proof that wave-like activity patterns return to their origins, like mathematical sine waves. The notion of reentry and adaptive resonance indeed suggests the opposite.

⁷Major thalamic divisions are called “nuclei,” but they are typically layered arrays of neurons that are tightly folded, so that they appear as clumps to the naked eye.

input and neuromuscular output use spatial arrays. The senses also converge on abstract spatial representations.

Thus the concept of a spatiotopic activity map (SA map) is useful, defined as a pattern of neuronal activity over a spatiotopic array of cells. We start with “activity maps” rather than structural maps, because only a subset of the cells in any spatiotopic array is active at any given moment. Therefore cortical and thalamic layers can be treated as four dimensional matrices with indices (i, j, k, t) , with positive or negative scalars for excitatory, or inhibitory cells at time t . Active neurons can then be labeled $\text{CELL}\# < i, j, k, t >$ and take on numerical values above or below a threshold during some brief time period, perhaps 0.1 s, since neurons in cortex may fire at a baseline 10 Hz rate. While this is highly simplified, it leads to a sensible conception of a brain-based GW.

Figure 2 can be viewed as SA maps that combine with others into SA streams; SA streams may compete or cooperate until some winner-take-all coalition results in a stable outcome, perhaps as briefly as the peak of the alpha or theta wave (Freeman, 2007). Together, the many active layers of the visual cortex can act as a resonant dynamical system that can yield an open set of visual experiences over time.

ANY-TO-ANY SIGNALING

Figure 4 shows the major long-distance connections in the C-T system of the macaque, including the basal ganglia (Modha and Singh, 2010). It is consistent with the notion of “any-to-any” signaling among SA arrays in cortex, mirrored in thalamic nuclei. Any cortical array can signal any other, both directly and via thalamic relays.

Modha and Singh also emphasize an “inner core” of high-density connections, consistent with a Neural Darwinist perspective of synaptic growth and myelination by use. In effect, the white matter pathways of the C-T system should reflect their traffic density, much as highway engineers would broaden highways to reflect their maximum usage, while allowing low-density access to residential streets. High-resolution tractography of the white matter pathways shows small-world network properties, consistent with efficient allocation of neuronal traffic flow (Achard et al., 2006).

In the visual cortex, dynamic SA coalitions may represent a single star on a dark night, a coffee cup on a nearby table, a simple event such as a cup falling off a table, and the textured surfaces of visual objects. Using the rich set of connectivities shown in **Figure 4**, such a stable visual coalition can signal to other C-T regions. The C-T system enables phase coupling and decoupling among active spatiotopic arrays (see also **Figures 5** and **6**). Steriade (2006) suggests a hierarchy of cortical oscillations ranging from <1 to 200 Hz, able to partly depolarize large numbers of neurons.

The functional anatomy of **Figure 4** suggests a capacity for “any-to-any” signaling among SA maps, streams, and coalitions. Other authors report that C-T connectivities can be described by small-world mathematics (Sporns and Honey, 2006), and Tononi’s formal measure phi provides a global index of connectivity in the C-T core. Such arrays of activity maps can also represent shifting attentional focus in the visual cortex (Itti and Koch, 2001).

SA coalitions follow trajectories that evolve by adaptive resonance. Edelman (1989) has made the case for genetic, ontogenetic,

and finally moment-to-moment selectionist processes in the brain. Each selectionist stage shows replication, variation, and selection, including the moment-to-moment dynamics of the C-T core. Dynamic activity travels along pathways and synapses built by use, which are constantly being pruned, beginning in the first month *in utero*.

Entire visuotopic arrays can play inhibitory or excitatory roles, so that SA maps can take on all positive or all negative values. This is useful when the visual system needs to focus on a single level of analysis, like a single colored dot in a Seurat painting, rather than a gestalt of many colored dots.

BROADCASTING: ANY-TO-MANY SIGNALING

A few ants can secrete alarm pheromones to alert a whole colony to danger, an example of any-to-many broadcasting among insects. In humans the best-known example is hippocampal-neocortical memory storage of memory traces in the neocortex by way of the hippocampal complex (Nadel et al., 2000; Ryan et al., 2001). Memories of conscious episodes are stored in millions of synaptic alterations in the neocortex (Moscovitch et al., 2005). We will return to this point in section “Voluntary Reports of Conscious Events.”

Computer users are familiar with global memory searches, which are used when specific searches fail. The C-T system may enable brain-based global memory searches. “Any-to-many” coding and retrieval can be used to store and access existing information (Nadel et al., 2000; Ryan et al., 2010). It is also useful for mobilizing existing automatisms to deal with novel problems.

Figure 4 shows how conscious visual stimuli may trigger a burst of any-to-many signaling from the right visual cortex to both hemispheres. By instructing subjects to attend to a lateralized right or left stimulus while keeping the eyes physically fixated on a target, either the right or left LED can be made conscious.

Notice that “any-to-many” signaling does not apply to the cerebellum, which lacks parallel-interactive connectivity, or to the basal ganglia, spinal cord, or peripheral ganglia. Crick and Koch have suggested that the claustrum may function as a GW underlying consciousness. However, the claustrum, amygdala, and other highly connected anatomical hubs seem to lack the high spatiotopic bandwidth of the major sensory and motor interfaces, as shown by the very high-resolution of minimal conscious stimuli in the major modalities. On the motor side there is extensive evidence for trainable voluntary control over single motor units and more recently, for voluntary control of single cortical neurons (Cerf et al., 2010). The massive anatomy and physiology of cortex can presumably support this kind of parallel-interactive bandwidth. Whether structures like the claustrum have that kind of bandwidth is doubtful. The sheer size and massive connectivity of the C-T system suggests the necessary signaling bandwidth for a human being to see a single near-threshold star on a dark night.

We do not know the full set of signaling mechanisms in the brain, and any current model must be considered provisional. Neural computations can be remarkably flexible, and is, to some degree, independent of specific cells and populations. John et al. (2001) has argued that active neuronal populations must have dynamic turnover to perform any single brain function, like active

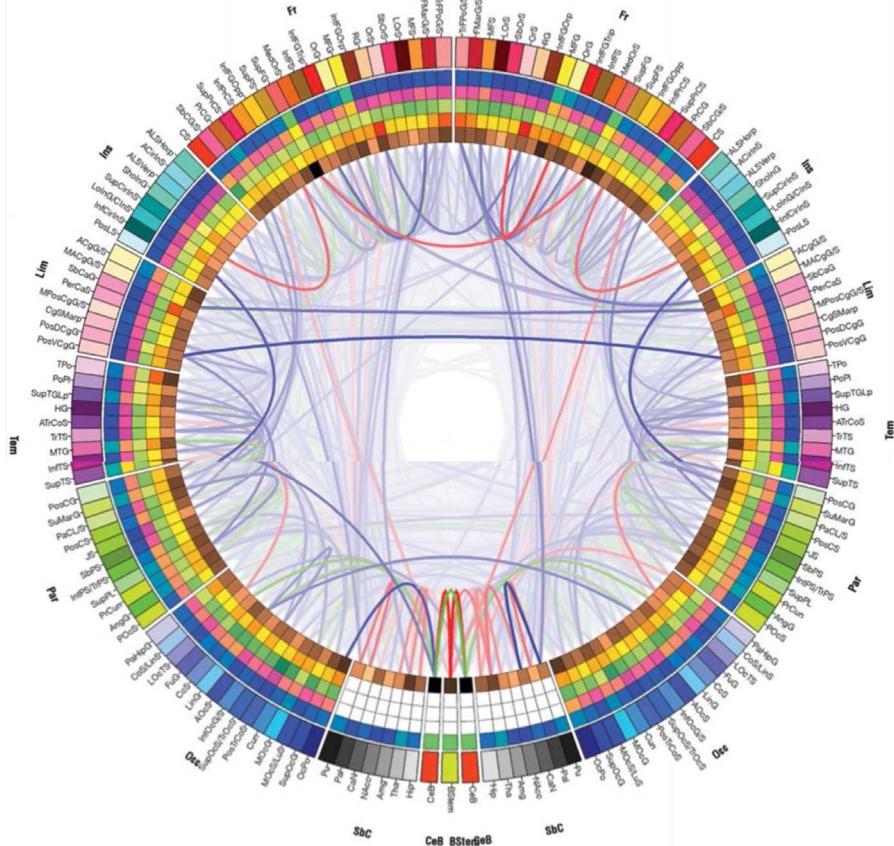


FIGURE 4 | The C-T complex supports any-to-any signaling. The cortico-thalamic system allows any spatiotopic activity (SA) array to signal any other. Combined with adaptive resonance, this allows an open set of cortical and thalamic coalitions to bind and broadcast information from any region to

any other. The left half represents the left hemisphere of the brain, whereas the right half represents the right hemisphere. The brain stem is shown at the bottom. Circular color bars at the bottom describe the scale of the corresponding anatomical ring.

muscle cells. Edelman and Tononi (2000) and others have made the same point with the concept of a dynamic core. GW capacity as defined here is not dependent upon the mere existence of anatomical hubs, which are extremely common. Rather, it depends upon a dynamical capacity, which operates flexibly over the C-T anatomy, a “functional hub,” so that activated arrays make up coherent “coalitions.”

The global neuronal workspace has been used to model a number of experimental phenomena. In a recent model, sensory stimuli mobilize excitatory neurons with long-range cortico-cortical axons, leading to the genesis of a global activity pattern among workspace neurons. This class of models is empirically linked to phenomena like visual backward masking and in attentional blindness (Dehaene and Changeux, 2005).

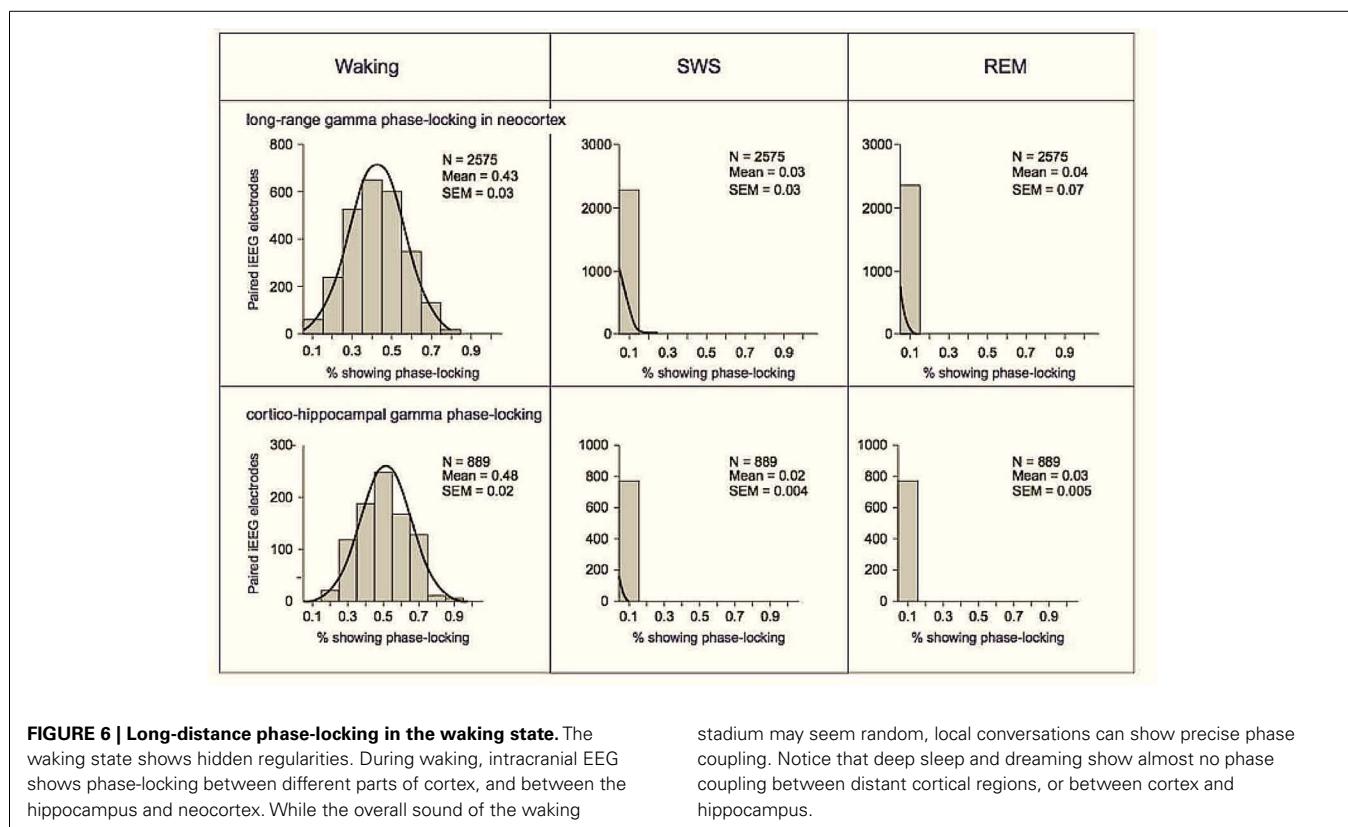
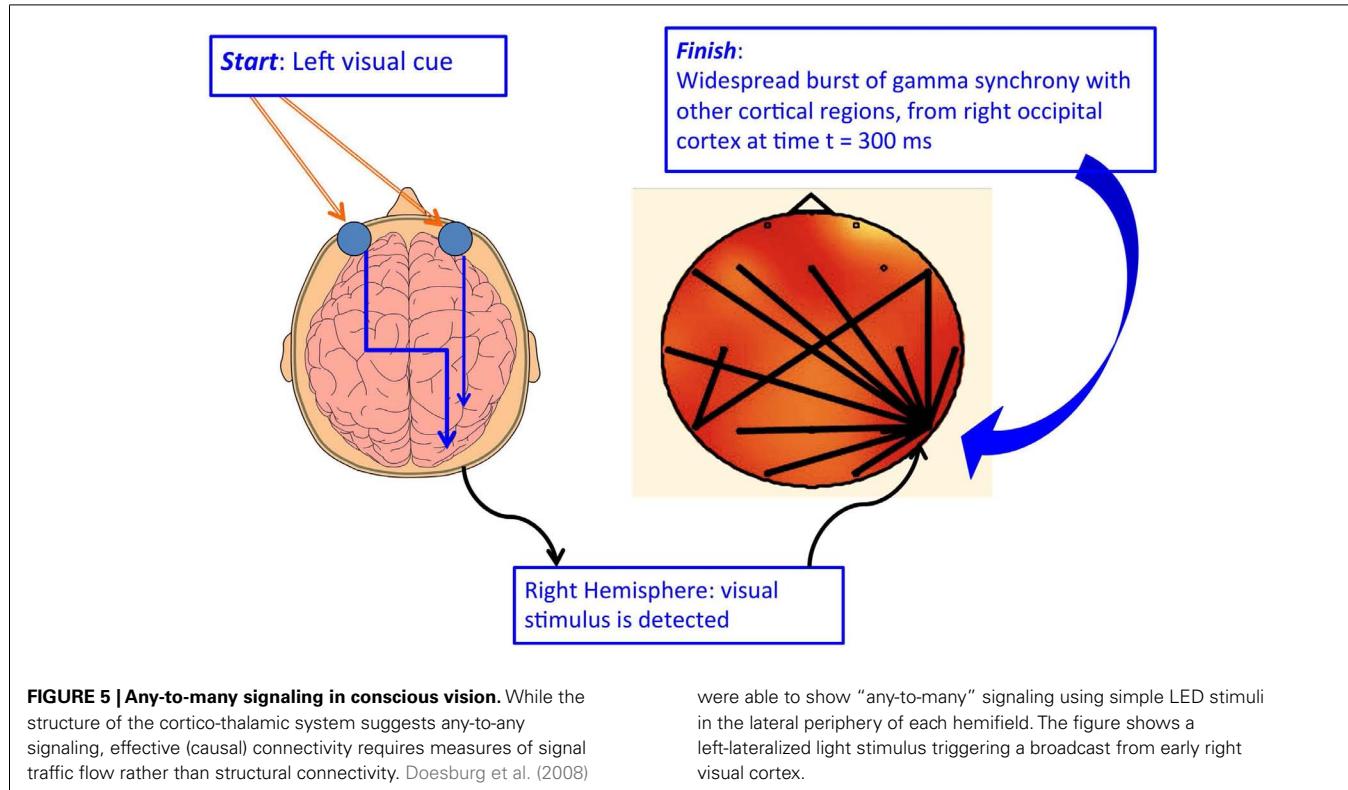
An advantage of the GW approach is that it is clearly functional; that is, it describes a known architecture for integrating and distributing focal information in a large-scale, massively parallel, brain-like organ. This functional architecture has been used for practical applications since Newell and colleagues developed it in the 1970s. However, it is a major challenge to scale neuronal models to the level of real brains. Current neuronal GWT models make use of simplified network architecture.

Franklin et al. (2012) have combined several types of computational methods using a quasi-neuronal activation-passing design. High-level conceptual models such as LIDA can provide insights into the processes implemented by the neural mechanisms underlying consciousness, without necessarily specifying the mechanisms themselves. Although it is difficult to derive experimentally testable predictions from large-scale architectures, this hybrid architecture approach is broadly consistent with the major empirical features discussed in this article. It predicts, for example, that consciousness may play a central role in the classic notion of cognitive working memory, selective attention, learning, and retrieval.

STATES AND CONTENTS

WAKING AND DEEP SLEEP

In mammals all goal-directed behavior occurs in the waking state, which enables fast, flexible, and widespread processing. State differences can be seen with the naked eye in the scalp EEG, which measures the post-synaptic activity of large numbers of cortical pyramidal neurons whose axons run orthogonal to the scalp. However, scalp recordings reflect only 0.1% of the cortical voltage, so that intracranial recordings give a much better picture.



Waking EEG looks irregular, complex, and low in amplitude, while deep sleep is marked by regular, simpler, and high-amplitude waves, reflecting coordinated “buzz-pause” firing among billions of neurons near 1 Hz. REM dreaming looks similar to waking, and dreamers often report rich conscious experiences.

Spontaneous conscious mentation occurs throughout the waking state, reflecting repetitive themes described as “current concerns”. Conscious mentation is also reported when subjects are awoken from REM dreams and even from slow-wave sleep. The last may reflect waking-like moments during the peaks of the delta wave (Valderrama et al., 2012).

Thus the waking state appears to enable brain-wide, adaptive resonance in the cortex and thalamus. By comparison, slow-wave sleep – the least conscious state of the daily cycle – may interrupt processing due to massive firing pauses in the trough of the delta wave (Edelman and Tononi, 2000).

Both sleep and waking are under precise homeostatic state control (John et al., 2001), controlled by thalamocortical circuits timed by basal brain nuclei. More than 200 types of epigenetic expression accompany sleep and waking in rats, suggesting many basic biological functions. Sleep appears to be a survival necessity, since rats die from complete deprivation after only 3 weeks. The exact reason is still unknown.

GLOBAL CHATTING, CHANTING, AND CHEERING

Global brain states can be compared to a football crowd with three states, “chatting,” “chanting,” and “cheering.”

Chatting describes the C-T activity of waking and REM dreams. It involves point-to-point conversations among spatial arrays in the C-T system, which can have very high S/N ratios, though they appear to be random when many of them take place at the same time. Like a football stadium with thousands of coordinated local conversations that are not coordinated globally, the average global activity is a low-level crowd roar, seemingly random, which appears to be fast and low in amplitude. Nevertheless, as we will see, direct cortical recordings show phase-coupled chatting in the C-T core appears to underlie specific cognitive tasks. Thus chatting activity gives the misleading appearance of randomness en masse, but it is in fact highly organized in a task-driven fashion. Because sports arenas show the same properties, the arena metaphor provides us with a useful reminder.

Chanting shows coordinated start-stop crowd activity, about once a second over a prolonged period of time, like the “buzz-pause” rhythm of billions of neurons in the C-T core, which results in global delta waves. Chanting sounds like chatting at the peak of the delta wave, followed by simultaneous pausing, which interrupts all conversations at the same time (Massimini et al., 2005). Breakdown of cortical effective connectivity during sleep.

Deep sleep (also called delta sleep or slow-wave sleep) is the least conscious state of the daily cycle, perhaps because 1–4 Hz pauses during the DOWN hemicycle of delta disrupt any cognitive tasks. However, there is some reported mentation when people are awoken from delta sleep, suggesting that the peak activities may be conscious or preconscious, so that during the process of awakening the chanting state is interpreted by the cortex in coherent ways. Because we cannot get instant reports from sleep, it is not obvious whether the reports are shaped by the seconds it takes to wake up.

Finally, a stadium crowd may cheer when a team scores a goal or makes an error. This corresponds to an “event-related” peak of activity. In the brain, the event-related potential (ERP) occurs when a significant or intense stimulus is processed, causing a stereotypical wave pattern to sweep through the brain. The event ERP shows a series of negative and positive deviations over a period of ~600 ms, corresponding to early sensory processing (N100), stimulus recognition, and attentional modulation (P200), decision making, and mismatch detection (P300 a and b), and stimulus meaning (N400 and longer). Revonsuo and colleagues have found that a conscious visual stimulus contributes a small negativity near the P3b wave (see Figure 8). ERP peaks and valleys are sensitive to numerous cognitive variables, but their overall shape tends to remain stable. ERP waves show both stimulus-triggered resetting and ongoing background activity.

In section “States and Contents” we suggested that the ERP reflects a repeating perception-action cycle that is mostly unconscious, but which embeds conscious moments lasting around 100 ms. The extended LIDA model interprets both conscious and non-conscious components of such “cognitive cycles” (Franklin et al., 2012).

Using direct brain recording in humans, Figure 6 shows that the waking state enables fast, long-distance phase-locking in the C-T core. Phase-locking is measured by lagged correlated activity among widely dispersed electrodes, and is believed to represent signaling among the cortical arrays being recorded. In this case the signal is measured in the low gamma band, 35–58 Hz, but strikingly similar findings are common from <1 to 200 Hz. Delta, theta, alpha, beta, and gamma phase coupling are commonly reported for a variety of tasks. This very wide band of apparently functional signaling is being studied intensively.

Phase coupling in the gamma range is tightly linked with the waking state. Slow-wave sleep and REM dreaming show a drastic drop in array-to-array gamma coherence. Observed gamma synchrony may reflect underlying theta-gamma cross-frequency coupling, which may in turn be carried by Slow Oscillation-theta-alpha coupling, as suggested by Steriade (2006). Slow Oscillations (SOs) continue throughout the daily cycle at <1 Hz. Widespread SO’s may change the firing threshold among billions of neurons, just as regional theta activity might do at a faster rate. A slow-to-fast wave hierarchy is one way to coordinate large populations of neuronal oscillators in the brain.

TASK-RELATED FREQUENCY COUPLING

A conscious task tends to recruit many cortical and subcortical populations. A range of evidence suggests that phase-coupled firing of large numbers of neurons is involved. Gamma and alpha-range synchrony are implicated in the conscious interpretation of ambiguous stimuli. However, alpha and theta activity in large populations of cells has been observed in other tasks as well. These ~10 Hz oscillations may serve to group faster rhythms such as beta and gamma. Both alpha-gamma and theta-gamma phase coupling have been observed.

Direct brain recording suggests that task-specific activity involves cross-frequency coupling at multiple spatial scales, linking, and unlinking multiple sites in the C-T core and its satellites. Single neurons can also self-adapt to regional theta oscillations

(Canolty et al., 2010; Canolty and Knight, 2012). Population oscillations and single neuronal spike timing may therefore mutually entrain.

Figure 7 shows direct cranial recordings from the left hemisphere in an epileptic patient, showing precise place-to-place signaling underlying the gross “random-looking” activity of the waking brain. Alpha-gamma and theta-gamma phase coupling serves to coordinate neurons that may be widely dispersed, but which support similar tasks.

In deep sleep, rapid interactive linking between regions may continue during the “up” sweep of the delta/Slow Oscillation, but it is inhibited during the “down” sweep (Edelman and Tononi, 2000).

CONSCIOUS CONTENTS

Traditional psychophysics compares conscious stimuli to each other, a method that dates back to Newton’s prism experiments. However, in a recent wave of research, conscious events are compared to unconscious ones that are physically similar, though they cannot be reported as conscious even under optimal conditions. Many experimental paradigms allow for such comparisons. We can therefore perform two kinds of comparisons for any conscious stimulus: psychophysical differentiation and iso-stimulus contours to define the feature boundaries of the percept and parametric variation to specify the boundaries between conscious and unconscious versions of the same event. When content parameters can be precisely controlled we can therefore make claims about

the contents of consciousness during waking, dreaming, twilight states, and even slow-wave sleep.

A CHANGING STREAM OF BOUND MOMENTS

The flow of conscious moments is often described as a stream (James, 1890). Under careful experimental conditions perceptual events appear stable and internally consistent. The Necker cube is a useful example of a bistable stimulus that is perceived as internally consistent at any given moment. Perhaps all the words of natural language are ambiguous in that they change meaning in different contexts (Baars, 1988). Under normal conditions only one meaning becomes conscious; the process of disambiguation is usually unconscious. The stream of consciousness is therefore constantly changing, but momentary events are perceived as stable and consistent. It is plausible that the stream’s contents emerge from ongoing winner-take-all competitions between activity streams (see Consciousness Enables Many Kinds of Access).

Here we focus on spatiotemporal coding as described above. However, this is a fast-moving frontier. While labeled line topographical resonance is a plausible contributor to the stream of specific conscious experiences, the same mechanism in the dorsal stream does not seem support conscious experiences directly (e.g., Milner and Goodale, 2008). This suggests that we may know some necessary but not sufficient conditions for conscious contents (viz., Baars, 1988). There are several other proposals for the ways in which conscious events may be encoded in cortical activity (Freeman, 2007). The full set of conditions is not yet known.

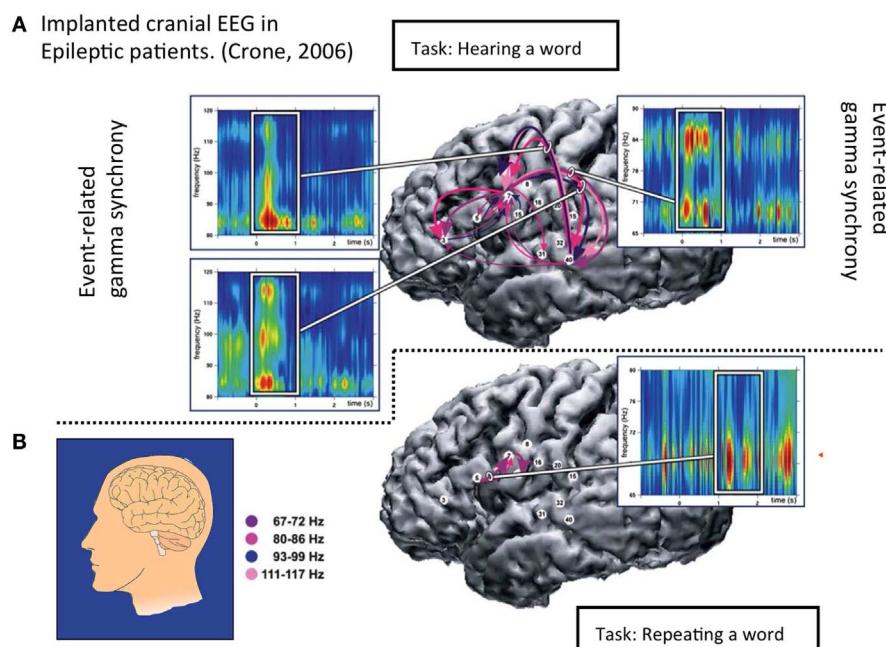


FIGURE 7 | Intracranial recordings in epileptic patients have 1,000 times the signal-to-noise ratio of scalp recordings, and therefore reveal much more detail. Crone et al. (2006) published these images of the left lateral hemisphere in a conscious epileptic patient before surgery, listening to a spoken word (**A**) and speaking it (**B**). Conscious patients in this procedure experience little pain under local anesthetic. White numbered disks are electrodes, and purple arrows between them indicate event-related synchrony (ERS) as shown in the graphs. In Task (**A**)

“Hearing a word,” ERS bursts in four gamma bands occur 100–600 ms post stimulus. In Task (**B**) “Speaking a word” ERS gamma starts before the response, and continue during the next two seconds in 100–200 ms bursts. Task (**A**) shows widespread left hemisphere ERS, while in Task (**B**) ERS is localized near Broca’s area for speech production. Other studies show precise ERS bursts in cortex for sensory processing, response organization and memory coding. Cortical synchrony may be a task-specific signaling code.

CONSCIOUS AND UNCONSCIOUS TASK ELEMENTS

Efficient cognitive tasks optimize tradeoffs between conscious and unconscious, and consciously mediated information processing. As described above, the ERP is a stereotypical waveform lasting about 600 ms, triggered by an intense or significant stimulus, and usually averaged over trials to allow common population responses to emerge while unrelated activity averages to zero. However, it has been difficult to separate the conscious and unconscious aspects of the ERP. New evidence shows that conscious (compared to unconscious) visual stimuli trigger a small difference in the visual ERP, emerging as a Visual Awareness Negativity (VAN) near the conventional P3b waveform, ~300 ms after stimulus onset and lasting for 100–200 ms (Gaillard et al., 2009). The VAN is believed to sweep forward from occipital-temporal regions (see **Figure 8**). These results have been replicated more than a dozen times, with different techniques. They suggest that the distinctively conscious signal adds only a small energy component to unconscious background processes, even though they may be triggered by the identical sensory event.

While this may seem surprising given the subjective richness of conscious events, it is not necessarily different from other kinds of signal processing. The most obvious example is radio waves, which use far more energy to propagate carrying waves than they

do for the contents of interest, which only modulate the carrier wave. While analogies must be treated with care, there is evidence that the intrinsic activity of the C-T core outweighs external input.

Thus cortical activity as shown in the visual ERP does not determine conscious contents directly. For example, the “novelty P3b” wave reflects population responses to unexpected events, which are processed unconsciously, but which may lead to conscious percepts. The P3 a and b components may therefore reflect preconscious processes. As suggested, the stereotypical ERP may represent a perception-cognition-action cycle in which a global broadcast “ignition” is only one brief component (see Sensory percepts vs. feelings of knowing below). **Figure 8** shows the consciousness-linked ERP component emerging in occipital cortex between 200 and 300 ms post stimulus, and followed by a late positivity (LP).

As pointed out, the C-T complex also engages subcortical satellites like the cerebellum and basal ganglia, which do not determine conscious contents directly.

WINNER-TAKE-ALL COALITIONS: MICROSTATES

Figure 9 shows evidence for stable microstates in the EEG for both the rabbit and humans, exhibiting rapidly changing phase after ~100–200 ms, the rate of theta oscillations (Freeman, 2007). Other laboratories, using quite different methods, have

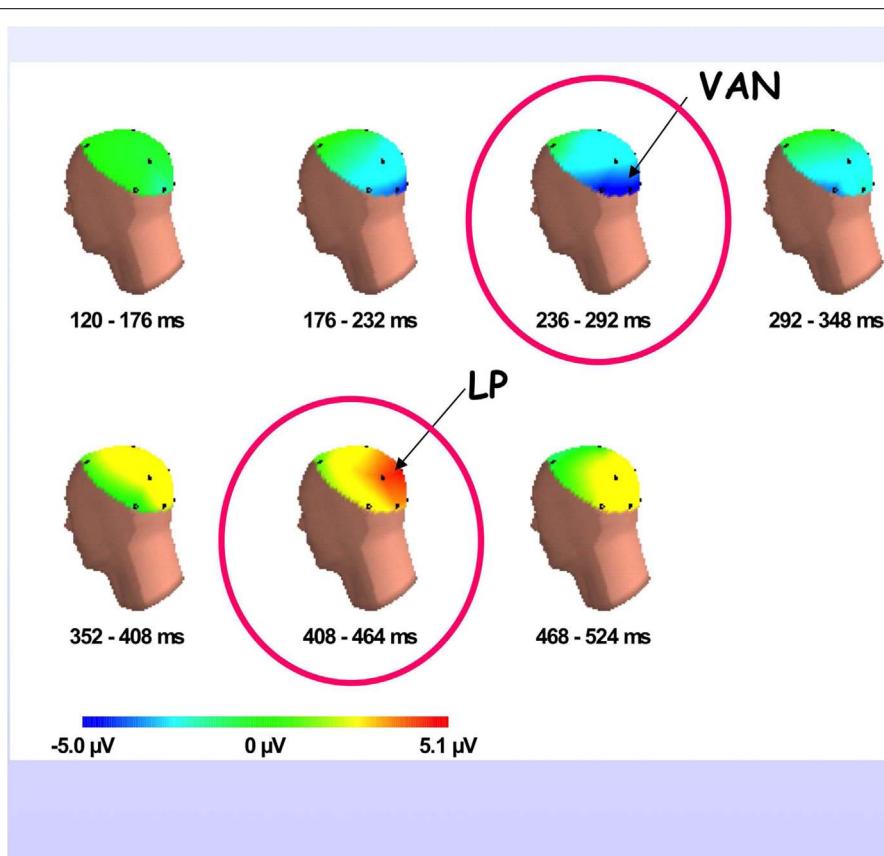


FIGURE 8 | VAN: the visual awareness negativity wave. The small conscious content-linked component of the visual awareness ERP (VAN) appears about 300 ms after a visual stimulus and lasts for 1–200 ms. It

appears to sweep forward, and may be followed by a late positivity (LP). The VAN has been observed in more than a dozen studies comparing conscious to unconscious visual stimuli that are identical at the retina.

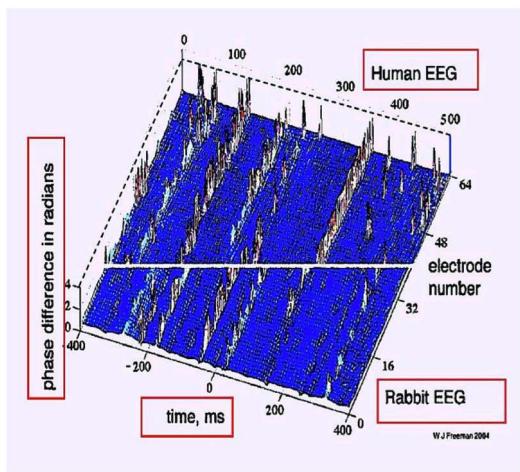


FIGURE 9 | EEG microstates at theta rates in humans and rabbits.

Global microstates have been found using several methods. Freeman et al. (2003) used Hilbert analysis to observe fast phase changes at theta rates in entire cortical hemispheres in humans and rabbits. The vertical axis shows phase differences in radians, revealing stable states for 100–200 ms, interrupted by a hemisphere-wide near-instantaneous collapse for 5–10 ms before a new equilibrium is achieved. The horizontal axes show time (in ms) and electrodes.

also reported momentarily stable, content-sensitive microstates. The current theory suggests that microstates represent binding and broadcasting equilibria involving dynamic coalitions of adaptively resonant populations of neurons. This view seems quite compatible with sophisticated theoretical work by Freeman and Kozma. Crick and Koch (2005) suggested the term “coalitions” for this general concept.

Coalitions can be linked to conscious (rather than unconscious) stimuli even at the level of single cell activity. In an important experimental series, Logothetis and colleagues recorded directly from visual cortical neurons in the macaque, while presenting rivaling stimuli known to evoke activity from edge detection to visual object identification. Neuronal competition and cooperation appeared in established cortical feature maps, with a decisive majority of neurons firing to the conscious percept only in the visual object area, corresponding to inferior temporal cortex (area IT) in humans. Thus, conscious, but not unconscious, visual stimuli evoked a decisive “vote” of neurons in IT.

Logothetis et al.’s research program continues to yield spectacular results including the strongest direct evidence for global distribution of conscious visual stimulation to date (Panagiotaropoulos et al., 2012). More recent studies indicate involvement of the dorsolateral prefrontal cortex, which also has visuotopic maps. However, the bilateral removal of the prefrontal cortex does not abolish binocular rivalry, suggesting, again, that binding of visual input is not localized in a single region, but varies dynamically depending on stimulus, task, anticipatory set, emotional valence, memory load, distractors, and perhaps cortical resource allocation.

Several laboratories have now reported EEG microstates lasting 100 ms or longer, and correlated with conscious contents. In

our terms, microstates show momentary equilibria in the C-T core. Freeman and colleagues provide a sophisticated non-linear dynamical interpretation of microstates. Adaptive resonance among topographically organized arrays leads to a dynamic system trajectory that may (1) land in a local minimum, (2) diverge and dissipate, or (3) even lead to a destructive positive feedback loop. Positive feedback among excitatory neurons resembles an epileptic seizure, and large-scale simulations of the cortex do show such activity. However, excitatory activity is constantly regulated by local, regional, and global inhibition, keeping the resonant trajectory between dysfunctional boundary states. Over time, adaptive resonance therefore either disappears, or it may evolve a temporary but stable equilibrium.

Since “binding” and “broadcasting” involve adaptive resonance, the distinctive type of signaling in the C-T core, dGW suggests “binding resonance” to define the winner-take-all gestalt that becomes conscious and “broadcasting resonance” to propagate the winning gestalt to receiving networks.

A human analogy for broadcasting resonance would be a college class equipped with feedback clickers to respond to points made by a lecturer. Signaling goes in both directions, and clickers may be used to ask the lecturer to explain some point in greater detail, so that signal receivers can influence the duration of the broadcast. Broadcasting leads to widespread adaptation among receiving processors (Baars, 1988, see Sensory Percepts vs. Feelings of Knowing); by analogy, a lecture leads to widespread learning. If students need more time to understand a broadcast, they can use their clickers to keep the broadcast signal flowing until they can adapt to it.

The most interesting possibility is that reported microstates reflect both binding and broadcasting in the brain, much like double-wing Lorenz attractors. Using a Hilbert analysis of the cortical EEG, Freeman et al. (2003) report that electrophysiological equilibria reflect a hemisphere-wide spatial phase change, occurring at theta rates. Future work will explore this link.

Recent evidence shows that single neurons can phase-tune their firing to regional theta oscillations, allowing single cells to coordinate with population activity (Canolty et al., 2010). Single cortical neurons can also come under voluntary control (Cerf et al., 2010), suggesting that even single units can serve executive goals.

CONSCIOUS MOMENTS ARE EMBEDDED IN CYCLES

The stereotypical ERP spins out over some 600 ms, reflecting basic population activities in processing incoming signals. Conscious moments are believed to last only 100–200 ms, and their distinctive signal appears to be small (Koivisto and Revonsuo, 2007). Franklin et al. (2012) have suggested that binding and global broadcasting is only one part of the recurrent action-perception cycle (Fuster, 2004). While conscious moments lasting ~100 ms may interfere with each other, the unconscious or automatic half-second of the cognitive cycle may overlap. Automatic skill components interfere less with each other than conscious contents.

Conscious binding and broadcasting seems to be largely a cortical-thalamic event that may trigger output to the striatum of the basal ganglia, a major output hub, which loops back to thalamus and cortex, providing a pathway for a recurrent unconscious cycle (Cole and Schneider, 2007).

SENSORY PERCEPTS VS. FEELINGS OF KNOWING

James (1890) emphasized that conscious contents are not just sensory; they include what he called “fringe” experiences or FOK, such as the TOT experience, accurate judgments, concepts (as opposed to percepts), relational terms like “and,” “or,” “but,” vivid expectations, beliefs, and other non-sensory but reportable events (**Table 1**). FOK show distinct BOLD activity in the frontal cortex, with fewer sensory maps than posterior cortex (Maril et al., 2001). Frontal cortex impairments also increase vulnerability to word retrieval difficulties (Pannu and Kaszniak, 2005).

Dynamic Global Workspace theory suggests that FOKs are bound and propagated from non-sensory regions of cortex, such as the classical association areas, frontoparietal regions, and the anterior temporal cortex. Brain imaging suggests that semantic knowledge is distributed in temporal, frontal, and parietal cortex, while sensory regions are recruited for imagery, inner speech, and motor activities that are associated with abstract concepts. However, effortful FOKs appear to be relatively localized in DL-Pfc and ACC regions. Effort-related BOLD activity spreads outward as tasks grow more difficult (**Tables 2** and **3**).

Feelings of knowing are not imprecise in their underlying contents. They are simply subjectively vaguer than the sight of a coffee cup (Baars, 1988), lacking clear figure-ground contrast, differentiated details and sharp temporal boundaries. However, concepts, judgments, and semantic knowledge can be complex, precise, and accurate. The history of science and mathematics is filled with examples of accurate, but subjectively vague, insights that could not be articulated and tested until much later.

Table 1 | Subjectively vague feelings of knowing.

The tip of the tongue	TOT state shows semantic knowledge while searching for a missing word. TOT's activate prefrontal regions fMRI (Maril et al., 2001*).
Judgments	Judgments of beauty, goodness, moral value, confidence, familiarity, rightness, etc., lack perceptual features and shapes, figure-ground contrast, specificity in space, time and content, etc. In contrast, conscious percepts are always particulars.
Relational terms	Prepositions like “to,” “from,” “by,” and relational terms like “above,” “below,” “before,” and “after” are not perceptual, though they may evoke feelings of knowing.

Table 2 | A summary of hypotheses.

Any-to-many binding and broadcasting in the C-T system
Perceptual consciousness vs. feelings of knowing
Dominant coalitions
Task-related chatting via phase coupling and decoupling
Feelings of effort interpreted
Voluntary control involves binding and broadcasting in the prefrontal lobe
Voluntary attention
The relationships to states of consciousness
The generality of adaptive resonance in C-T biocomputation

Thus, Fermat's Last Theorem occupied mathematicians for centuries, as a compelling but unproven FOK. Inexplicit but subjectively persuasive FOKs are a recognized route to mathematical discovery. In the brain, semantic knowledge is widely distributed and seems to lack figure-ground contrast.

Tip-of-the-tongue experiences can be induced by asking for the technical names of familiar facts. The question “What are two names for flying dinosaurs?” may elicit strong FOK. Subjects who cannot recall those names still choose accurately and quickly between “pterodactyl” and “brontosaurus.” Semantic knowledge may be fully primed in TOT states, before the lexical form of the missing words can be recalled. Such FOK commonly occur when we have compelling and accurate expectations and intentions. They are not limited to language.

Psycholinguistic experiments have long shown that semantic access occurs routinely without lexical recall. Thus many meanings do not have to be expressed in words. FOK can therefore be conceptually rich and precise (Baars, 1988).

Table 3 | Some testable predictions.

1. The cortico-thalamic system supports any-to-many binding and broadcasting of conscious contents. A bound conscious gestalt may emerge from anywhere in the cerebrum, and spread globally to all other regions for ~100 ms.
2. Receiving networks adapt to novel information from broadcast sources. After widespread receiver adaptation (updating), broadcasts are driven out of consciousness by competing inputs.
3. Posterior cortex generates perceptual conscious contents, while explicit feelings of knowing are generated from non-sensory cortex. (Frontal, anterior temporal, parietal). Many cognitive tasks involve both a conscious perceptual and a reportable semantic broadcast, as shown in the extended ERP.
4. Since nearly all cortico-thalamic links are bidirectional, the cerebrum supports very widespread adaptive resonance (reentrant signaling). Signaling of conscious contents is superimposed on baseline resonant activity in the C-T core. Because of spatiotopic array organization in the cortex and thalamic nuclei, content signaling in the cerebrum is simultaneously spatial and temporal.
5. Goal-directed signaling in the C-T core is waking state dependent. Waking, dreaming and slow-wave sleep reflect distinct global modes. However, even slow-wave sleep may support waking-like activity during the UP phase of the slow oscillation.
6. While many spatiotemporal codes may exist, cross-frequency phase coupling is thought to integrate the full range of C-T rhythms. Because conscious sensory events are integrated within 100 ms periods, 4–12 Hz rhythms may underlie conscious moments.
7. Effortful voluntary control involves binding and broadcasting from frontoparietal regions. Mental effort is an FOK that is associated with major cognitive styles like persistence and general intelligence.
8. The hippocampal complex supports conscious event perception, as well as serving to encode episodic memory traces in multiple brain regions. Hippocampal lesions often lead to cortical reorganization of conscious sensory functions.

MENTAL EFFORT

The sense of mental effort is an important FOK. It commonly occurs in the TOT experience, but also in a range of other tasks that are difficult, slower than expected, require persistence, or involve goal conflict. Duncan and Owen (2000) have done extensive research on this topic, and found that subjective mental effort activates executive regions of the frontal cortex (DL-PFC and ACC) across a variety of tasks. Mental effort may underlie the general intelligence factor *g*, which is found robustly across tests, tasks, and cultures. It may also correspond to the major personality variable persistence, which may explain its cross-cultural relevance, since persistence is needed for high performance in tasks that may be desirable in different cultural contexts. Baumeister has discovered closely related phenomena under the labels “willpower” and “ego depletion.” While these terms have an old-fashioned flavor, they are solidly based in evidence. For example, subjective mental fatigue accumulates during the waking day and may be relieved by increasing self-esteem.

Duncan and Owen (2000) explain subjective effort by a Multiple Demand System (MDS) of the prefrontal cortex. MDS is closely related to dGW, since any region of the C-T system may serve as a binding and broadcasting site according to the argument we have developed.

VOLUNTARY ATTENTION

In a GW perspective, attention is any process that facilitates access to conscious contents, much as eye movements enable access to specific visual events. Visual attention is associated with the eye movement control system, the frontal eye fields and superior colliculi. Cole et al. (2012) have proposed a “cognitive control system” for voluntary attention that involves global connectivity from the prefrontal cortex, and which predicts cognitive control and intelligence. This seems to be another aspect of the mental effort phenomenon discussed above. Goal-driven attention also selects conscious sensory information to recruit, plan, and carry out extended tasks in the world.

The key claim from GW theory regarding dynamics is therefore that conscious contents bind sensory or other cortical-thalamic contents into a winner-take-all coalition, able to broadcast to other regions of the C-T system. No single C-T location is a privileged source.

Figures 1 and 10 (above) showed four hypothetical “binding and broadcasting” sources in cortex, V1 for simple stimuli like a bright star on a dark night, a set of visual areas including IT for object representation at multiple levels, and the MTL/rhinal cortex for event organization.

Prefrontal activation across multiple tasks demanding mental effort (Duncan and Owen, 2000) suggests that sensory conscious experiences are bound and broadcast from the classical sensory regions in the posterior cortex, while voluntary effort, reportable intentions, feelings of effort, and the like, have a prefrontal origin, consistent with brain imaging findings. **Figure 10** also suggests that FOK and effort are bound and broadcast from the prefrontal cortex, notably the dorsolateral and anterior cingulate regions, areas that are known to be involved in subjective feelings of effort.

These findings suggest an hypothesis about sensory consciousness compared to “fringe” FOK, feelings of effort, and reportable

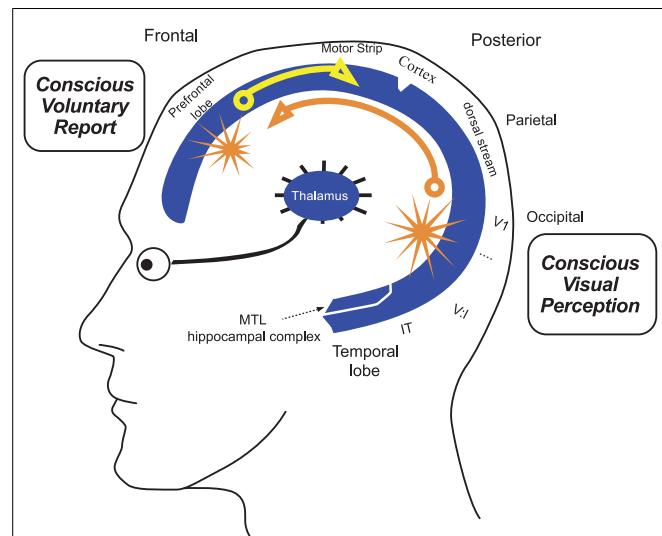


FIGURE 10 | Perceptual experiences vs. feelings of knowing (FOKs). This dGW cartoon shows an occipital broadcast (which must mobilize parietal egocentric and allocentric maps as well) evoking spatiotopic activity in the prefrontal cortex, which is known to initiate voluntary actions (see yellow arrow). Prefrontal activity is shown as a second global workspace burst, consistent with **Figure 1**. Visuotopic coding is preserved in the dorsolateral prefrontal cortex. The ability to voluntarily report or act upon a spatially specified stimulus follows from this double binding and broadcasting event. Thus, a posterofrontal broadcast is quickly followed by a centrifugal burst from prefrontal regions, including supplementary motor, premotor, and motor cortex. Posterior binding and broadcasting is experienced as a visual event framed in nearby space, while the prefrontal broadcast is a feeling of knowing (FOK) or “fringe” experience (James, 1890).

voluntary decisions. These reportable but “vague” events have been discussed since William James, who gave them equal importance to perceptual consciousness. fMRI studies show that they predominantly involve prefrontal regions, even across tasks that seem very different.

Because of the small-world connectivity of white matter tracts, different integration and distribution hubs may generate different global wave fronts. The sight of a coffee cup may involve an inferotemporal hub signaling to other regions, while the perception of music may emerge from Heschel’s gyrus and related regions. Reportable experiences of cognitive effort might spread outward from a combined DL-Pfc/ACC hub.

CONSCIOUS EVENTS EVOKE WIDESPREAD ADAPTATION OR UPDATING

What is the use of binding and broadcasting in the C-T system? One function is to update numerous brain systems to keep up with the fleeting present. GW theory suggested that consciousness is required for non-trivial learning (i.e., learning that involves novelty or significance) (Baars, 1988). While there are constant efforts to demonstrate robust unconscious learning, after six decades of subliminal vision research there is still surprisingly little proof. Subliminal perception may work with known chunks, like facial expressions. While single-word subliminal priming appears to work, Baars (1988) questioned whether novel two-word primes would work subliminally. The subliminal word pair “big house” might prime the word “tall,” while “big baby” might not, because it

takes conscious thought to imagine a baby big enough to be called tall. In general, the more novelty is presented, the more conscious exposure is required.

It follows that dGW should predict widespread adaptive changes after conscious exposure to an event. That is indeed the consensus for hippocampal-neocortical memory coding (Nadel et al., 2012). However, the hippocampal complex is not currently believed to enable conscious experiences. Nevertheless, episodic memory is by definition “memory for conscious events” (See The Hippocampus and Conscious Contents: a Novel Prediction). Conscious events trigger wide adaptation throughout the C-T system, and in subcortical regions that are influenced by the C-T system. **Figure 11** shows how novel, conscious tasks trigger widespread adaptive processing in the C-T core, while the identical task after repetition to the point of automaticity no longer requires high-metabolic C-T processing. Episodic, semantic, and skill (procedural) processing all follow the same curve of high-metabolic processing to novel, conscious learning followed by a drastic drop in conscious access and metabolic (BOLD) activity after learning.

Figure 11 shows an example for auditory search, showing BOLD peaks in many cortical locations, but fading after automaticity due to practice, except in primary auditory cortex. Performance improves as the BOLD signal fades and conscious access recedes (Schneider, 2009). Cortical BOLD fading after training is a robust fact, indicating that conscious (reportable) events evoke widespread adaptation at multiple levels, from single neurons to entire brains (Baars, 1988; Gomez et al., 2009). Adaptation to novelty has been proposed to be one of a small set of necessary conditions for conscious experience (Baars, 1988, Chapter 12).

CONSCIOUSLY MEDIATED COGNITION

Most cognitive tasks combine conscious and unconscious elements. If we assume from a GWT perspective that conscious task moments allow novel integrations while unconscious ones

implement automatic routines, we can see why fill-in sentences, cued recall, remote associates, “Aha!” experiences and ambiguous stimuli are consciously mediated. Baars and Franklin (2003) suggest that all working memory tasks are consciously mediated – inner rehearsal, executive control, and other WM elements are all claimed to require a conscious cue to execute. If we rehearse seven numbers, some are conscious at any moment, but others are not. We are not aware of non-rehearsed items at any moment in a working memory task, nor of the important role of the basal ganglia in controlling inner speech, or of the automatic (habitual) components of any task. Completely conscious tasks are rare.

Standard tasks in cognitive neuroscience are generally consciously mediated, yet classical notions of cognition give no functional account of that easily observable fact. Standard cognitive tasks like working memory recall could not be done if they were masked by momentary distractions. Apparently the conscious elements of cognitive tasks are often necessary for those tasks to be carried out. Conscious mediation of task elements appears to be mandatory, not optional.

From a GW point of view the answer is straightforward: all tasks that involve unpredictable choice points require global broadcasting to recruit knowledge sources. Baars (1988) devoted several chapters to the near-universality of ambiguity in human cognition, including perception, working memory, memory retrieval, voluntary attention, action planning, and voluntary motor control. When fully predictable tasks are rehearsed to the point of automaticity they tend to drop from consciousness; conversely, when ambiguity is introduced in predictable situations, more conscious involvement is required. The first GW architecture developed by Newell and coworkers was explicitly designed to resolve the multi-level ambiguities of auditory words spoken in a normal noisy space.

Thus GW theory provides a general answer to the question: why are consciously mediated tasks so common? At the same time, it

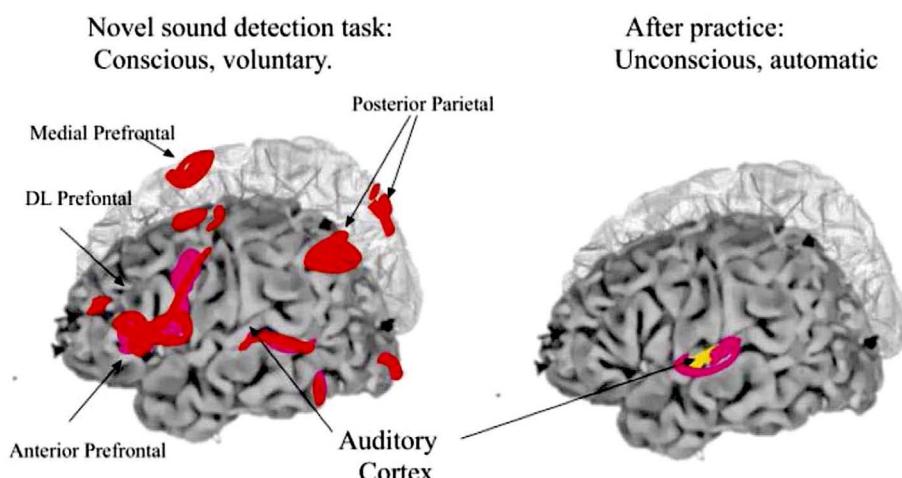


FIGURE 11 | Cortical adaptation as a novel task becomes automatic.

Schneider (2009) summarized a large literature on habituation of over learned skills. On the left, high BOLD activations are shown in the cognitive control system of the cortex, including the posterior parietal cortex and executive regions of the prefrontal cortex. After the task is practiced to the point of

automaticity, only the auditory region of the temporal cortex shows fMRI activity, as required for the auditory detection task. The BOLD signal corresponds to neuronal population activities. These results are consistent with dynamic Global Workspace theory in that the novelty stage involves more conscious and voluntary processing.

suggests a qualitative answer to the close link between task novelty and conscious access.

VOLUNTARY REPORTS OF CONSCIOUS EVENTS

Conscious contents are routinely assessed by voluntary report, as we know from 200 years of scientific psychophysics. Yet the reason for that fact is far from obvious. Any theory of consciousness must ultimately explain the basic fact that we can voluntarily report an endless range of conscious contents, using an endless range of voluntary actions. GWT suggests the shape of an answer.

Voluntary control is one kind of consciously mediated process. As we learn to ride a bicycle for the first time, each movement seems to come to consciousness. After learning, conscious access drops even as BOLD activity in the C-T core declines (**Figure 11**). We postulate that conscious involvement is necessary for non-trivial acquisition of knowledge and skills, and that the period of conscious access enables permanent memory traces to be established.

While “verbal report” is the traditional phrase, reports do not have to be verbal – any voluntary response will work. Broca’s aphasics who are cannot speak can point to objects instead. Locked-in (paralyzed) patients, who seem to be comatose, can learn to communicate by voluntary eye movements. Thus “verbal report” should be called “accurate, voluntary report,” using any controllable response.

Voluntary actions can point to objects and events. A “match to sample” task is commonly used to indicate the similarity of two conscious events, and to specify just noticeable differences. Pointing occurs naturally when mammals orient to a novel or significant stimulus. Children develop pointing abilities using “shared attention” in early childhood.

If we assume for simplicity that conscious contents emerge in posterior cortex, and that voluntary actions emerge in frontal and parietal cortex, we can ask the question in dGW terms: how is it that a posterior “binding and broadcasting” event is transformed into a frontally controlled action?

These facts raise the question of how accurate signal transmission occurs between sensory arrays and frontal executive control. In the case of pointing to a single star on a dark night, the physical minimum of light quanta in the retina can be amplified and transmitted to prefrontal cortex, which can control the movement of a single finger to point to the star. Even more remarkably, single neurons in the temporal cortex have been shown to be fired at will in surgical patients using intracranial electrodes, providing only that conscious sensory feedback is given during training (Cerf et al., 2010). Thus the physical minimum to the eye can accurately translate into “any” voluntarily controlled single cell, used as a sensory pointer. Given a million foveal cells for input, and perhaps billions of cortical cells for output, “any-to-any” mapping in the brain can involve remarkably large numbers. With accurate psychophysical performance in both tasks, the signal to noise ratio from receptor to effector cell can approach the physical limit. This precision needs explanation in terms of conscious input and voluntary control.

Figure 10 gives a dGW account of a conscious sensory experience spreading proposes a double broadcast, the sensory one

emerging from occipitoparietal regions, and the voluntary one (which may be subjectively effortful) coming from prefrontal executive cortex.

This also suggests an explanation for the standard index of voluntary report. When we report a star on a dark night, posterior broadcasting may lead to frontal binding and ultimately a frontal broadcast. Frontoparietal regents are driven by posterior sensory projections when they become conscious. Because of the striking similarities of spatiotopic coding in frontal and posterior cortices, we can image that sensory consciousness can also trigger a new binding, and broadcast an event in the frontal cortex. Voluntary action is therefore an extension of GW dynamics.

Figure 12 shows an intracranial electrode grid recording of the left hemisphere of an epileptic patient with implanted cortical electrodes for presurgical testing. It shows a word being perceived consciously, and after a suitable delay, being reported. The dependent variable in **Figure 12** is gamma synchrony bursts in four different frequency ranges (67–117 Hz).

For the first time in history we are able to look at the cortex in real time and with high spatial resolution as the patient hears a spoken word and then reports it. In both conditions, hearing and reporting, we can see phase coupling bursts in the left LIFG (lateral inferior frontal gyrus), Broca’s area. A dGW interpretation would suggest a broadcast from auditory cortex in the first condition (a), spreading to Broca’s area and setting up adaptive resonance lasting for several seconds (about the duration of working memory, 10–30 s). When the patient repeats the word in part (b) the relevant speech production regions are already primed by the stimulus broadcast. This account is still sketchy, since we only see two snapshots in time. However, we can use dGW to make more precise and testable predictions.

THE HIPPOCAMPUS AND CONSCIOUS CONTENTS: A NOVEL PREDICTION

Hippocampus is generally thought to encode episodic memories but not conscious experiences. This is puzzling, since episodic memories come from conscious episodes in the first place. It would be odd if these twin functions were not linked.

Because the neocortical-thalamic system arose with mammals some 200 million years ago, some kinds of consciousness appear to be evolutionarily ancient. Neurobiologists are increasingly searching for brain homologs and analogs of the C-T system found in mammals. What G. M. Edelman calls “higher level” (symbolic) consciousness would be based on the frontal expansion of the primate brain, giving rise to novel symbolic and language-dependent capacities.

In humans there is no anatomical or physiological barrier between paleocortex and neocortex. We therefore propose, as an hypothesis, that like neocortex, the hippocampal complex helps to specify conscious experiences as well as encoding episodic memories. dGW proposes that any region of neocortex can function as a source or receiving region for conscious experiences. We would therefore add “paleocortex,” including the hippocampus, to “neocortex.” What may be distinctive about the hippocampal complex is its role in encoding and decoding of conscious episodes into widely distributed memory traces. There is direct evidence from brain recordings in the MTL (Panagiotaropoulos

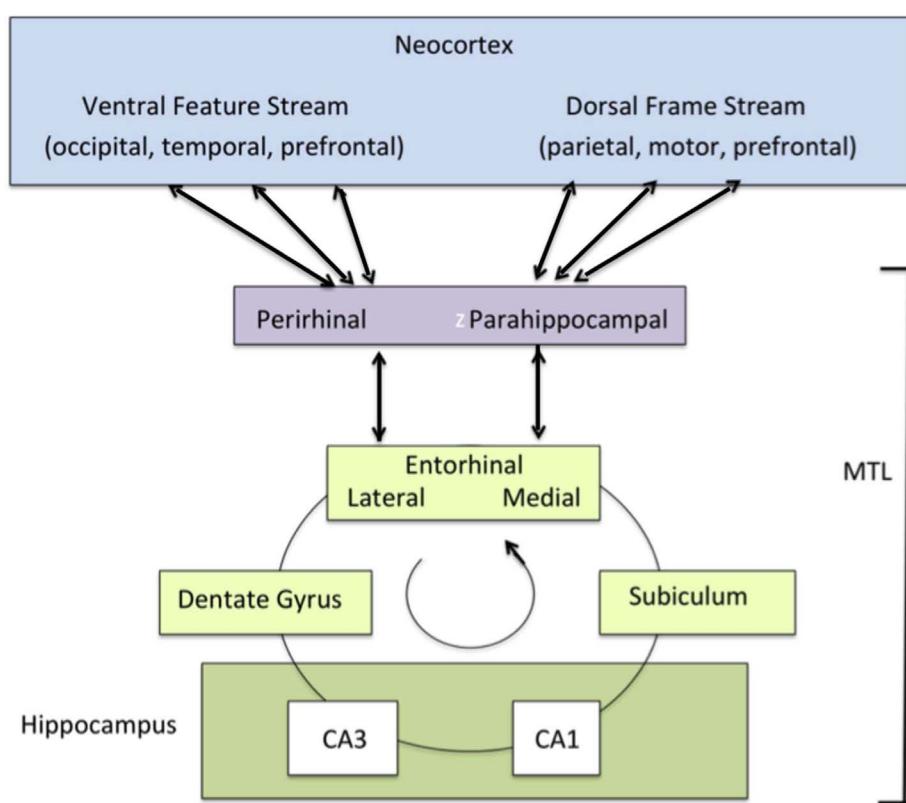


FIGURE 12 | Hippocampal-neocortical binding and broadcasting. The hippocampal complex in relation to the neocortex. Notice that the flow of information streams in both directions, with the interesting exception of the subiculum. Thus resonant adaptation can take place in almost all regions of the hippocampal-neocortical complex. This region appears as a major

convergence and broadcasting zone, and has been proposed to combine the dorsal and ventral streams of visual cortex (Shimamura, 2010). Direct MTL neuronal recording shows responding to conscious, but not unconscious visual input. Phylogenetically, the hippocampus is ancestral to the neocortex, and controls a fully autonomous sensorimotor brain.

et al., 2012), which follow the conscious percept in a binocular rivalry task, rather than the physical stimulus.

This seems to be contradicted by the classical case of patient HM, a rare case of very precise surgical excision of the bilateral MTL, which contains the hippocampal complex. HM appeared to have normal conscious perception, while his episodic memory was lost. HM's surgery therefore seemed to dissociate conscious perception from memory. HM was studied over a 60-year period and is still an important source of insight.

Cortical impairments are often difficult to interpret functionally, because patients spontaneously reorganize their behavior and cortex to compensate. For example, cutting the corpus callosum destroys about 200 million axons linking the hemispheres, but patients learn to use cross-body arm and eye movements to compensate. Early split brain patients seemed to show no impairments at all; careful experimental tests later on showed the precise nature of their functional losses. A similar debate occurred about the results of prefrontal lobotomy, stroke, minimal brain damage, and the like.

SUMMARY

A GW is a functional hub of signal binding and propagation in a population of loosely coupled signaling agents. Neurons and

neuronal cell assemblies can be defined as such agents when they respond selectively to input.

Conscious experiences may reflect a GW function in the brain. The brain has many anatomical hubs, but conscious percepts are unitary and internally consistent at any given moment. This suggests that a brain-based GW capacity cannot be limited to only one anatomical hub. Rather, it should be sought in a dynamic and coherent binding capacity – a *functional hub* – for neural signaling over multiple networks. A number of findings are consistent with the theory.

In humans the C-T complex underlies reportable conscious percepts, concepts, FOK, visual images and executive functions. While subcortical areas are often sometimes proposed to specify conscious contents, the human evidence is slight and disputed. Because cortex and thalamus are interleaved so densely as to constitute a single functional system, we will refer here to the C-T system as a whole. C-T pathways permit constant reentrant signaling, so that multiple spatiotopic maps can sustain or inhibit each other. The daily states of the core are controlled by basal brain nuclei.

Global workspace theory follows the historic distinction between the “focus” of experience vs. the largely implicit background of experience. Extensive evidence shows that visual and

auditory consciousness flows from the respective sensory cortices to frontoparietal regions. This directionality differentiates GW dynamics from information integration theory and dynamic core theory.

Cortico-thalamic core is a great mosaic of multi-layered two-dimensional neuronal arrays. Each array of cell bodies and neurites projects to others in topographically systematic ways. Since all C-T pathways are bidirectional, signaling is “adaptively resonant” (reentrant). In this complex, layered two-dimensional arrays are systematically mirrored between cortex and thalamus, region by region.

The C-T nexus appears to be the most parallel-interactive structure in the brain, allowing for efficient signal routing from any neuronal array to any other. This connectivity is different from other structures that do not directly enable conscious contents, like the cerebellum. The cerebellum is organized in modular clusters that can run independently of each other, in true parallel fashion. But in the C-T core any layered array of cortical or thalamic tissue can interact with any other, more like the world-wide web than a server farm.

Cortico-thalamic pathways run in all canonical directions and follow small-world organization, so that each array is efficiently

linked to many others. The entire system acts as an oscillatory medium, with markedly different global regimes in conscious and unconscious states.

Global workspace dynamics interprets the traditional distinction between the “object” and “ground” of experiences as a directional flow between the moment-to-moment focus of conscious experience vs. the implicit background and sequelae of focal contents. The proposed directionality of broadcasting suggests a testable distinction with information integration theory and dynamic core theory.

For example, while sensory experiences are proposed to bind and broadcast from posterior cortex, “fringe conscious” FOK plausibly emerge from non-sensory cortices, and may therefore broadcast rostrocaudally.

Finally, dGW suggests the counter-intuitive idea that in intact humans, the MTL supports conscious episodes. The classic lesion case of HM seems to contradict that idea, but traditional psychophysical testing does not show episodic organization. We therefore suggest that the positive evidence from brain recording methods for conscious experience in MTL may outweigh lesion evidence against it, given the well-established tendency toward cortical reorganization after injury.

REFERENCES

- Achard, S., Salvador, R., Whitcher, B., Suckling, J., and Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *J. Neurosci.* 26, 63–72.
- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press. [Kindle edition, 2012].
- Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci. (Regul. Ed.)* 6, 47–52.
- Baars, B. J., and Franklin, S. (2003). How conscious experience and working memory interact. *Trends Cogn. Sci. (Regul. Ed.)* 7, 166–172.
- Baars, B. J., and Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: global workspace theory and IDA. *Neural Netw.* 20, 955–961.
- Bizley, J. K., Shinn-Cunningham, B. G., and Lee, A. K. (2012). Nothing is irrelevant in a noisy world: sensory illusions reveal obligatory within- and across-modality integration. *J. Neurosci.* 32, 13402–13410.
- Blumenfeld, H. (2012). Impaired consciousness in epilepsy. *Lancet Neurol.* 11, 814–826.
- Canolty, R. T., Karunesh, G., Kennerley, S., Cadieu, C., Koepsell, K., Wallis, J., et al. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17356–17361.
- Canolty, R. T., and Knight, R. T. (2012). The functional role of cross-frequency coupling. *Trends Cogn. Sci.* 14, 506–515.
- Cerf, M., Thiruvengadam, N., Mormann, F., Kraskov, A., Quiroga, R. Q., Koch, C., et al. (2010). Online, voluntary control of human temporal lobe neurons. *Nature* 467, 1104–1108.
- Cole, M. W., and Schneider, W. (2007). The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage* 37, 343–360.
- Cole, M. W., Yarkoni, T., Repovš, G., Anticevic, A., and Braver, T. S. (2012). Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *J. Neurosci.* 32, 8988–8999.
- Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Sem. Neurosci.* 2, 263–275.
- Crick, F. C., and Koch, C. (2005). What is the function of the claustrum? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1271–1279.
- Crone, N. E., Sinai, A., and Korzeniewska, A. (2006). High-frequency gamma oscillations and human brain mapping with electrocorticography. *Prog. Brain Res.* 159, 275–295.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: a systems level proposal for the neural substrates of recall and recognition. *Cognition* 33, 25–32.
- Dehaene, S., and Changeux, J. P. (2005). Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. *PLoS Biol.* 3:e141. doi:10.1371/journal.pbio.0030141
- Dehaene, S., and Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron* 70, 200–227.
- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.* 10, 204–211.
- Dehaene, S., and Naccache, L. (2001). Towards a cognitive permanence of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37.
- Demertzi, A., Soddu, A., and Laureys, S. (2012). Consciousness supporting networks. *Curr. Opin. Neurobiol.* pii:S0959–4388(12)00192–4.
- Doesburg, S. M., Green, J. J., McDonald, J. J., and Ward, L. M. (2009). Rhythms of consciousness: binocular rivalry reveals large-scale oscillatory network dynamics mediating visual perception. *PLoS ONE* 4:e6142. doi:10.1371/journal.pone.0006142.
- Doesburg, S. M., Roggeveen, A. B., Kitajo, K., and Ward, L. M. (2008). Large-scale gamma-band phase synchronization and selective attention. *Cereb. Cortex* 18, 386–396.
- Duncan, J., and Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483.
- Edelman, G. M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books Inc.
- Edelman, G. M., Gally, J. A., and Baars, B. J. (2011). Biology of consciousness. *Front. Psychol.* 2:4. doi:10.3389/fpsyg.2011.00004
- Edelman, G. M., and Tononi, G. (2000). *A Universe of Consciousness: How Matter Becomes Imagination*. New York: Basic Books Inc.
- Franklin, S., Strain, S., Snaider, J., McCall, R., and Faghhihi, U. (2012). Global workspace theory, its LIDA model and the underlying neuroscience. *Biol. Insp. Cogn. Arch.* 1, 32–43.
- Freeman, W. J. (2007). Indirect biological measures of consciousness from field studies of brains as dynamical systems. *Neural Netw.* 20, 1021–1031.
- Freeman, W. J. (2009). Vortices in brain activity: their mechanism and significance for perception. *Neural Netw.* 22, 491–501.
- Freeman, W. J., Burke, B. C., and Holmes, M. D. (2003). Aperiodic phase resetting in scalp EEG of beta-gamma oscillations by state transitions at alpha-theta rates. *Hum. Brain Mapp.* 19, 248–272.
- Fuster, J. M. (2004). Upper processing stages of the perception-action cycle. *Trends Cogn. Sci.* 8, 143–145.

- Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., and Hasboun, D. (2009). Converging intracranial markers of conscious access. *PLoS Biol.* 7:e1000061. doi:10.1371/journal.pbio.1000061
- Gomez, A., Rousset, S., and Baciu, M. (2009). Egocentric-updating during navigation facilitates episodic memory retrieval. *Acta Psychol. (Amst.)* 132, 221–227.
- Goodale, M. A., and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Itti, L., and Koch, C. (2001). Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2, 194–203.
- Izhikevich, E. M., and Edelman, G. M. (2008). Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3593–3598.
- James, W. (1890). *The Principles of Psychology*. New York: Holt.
- John, E. R., Prichep, L. S., Kox, W., Valdés-Sosa, P., Bosch-Bayard, J., Aubert, E., et al. (2001). Invariant reversible QEEG effects of anesthetics. *Conscious. Cogn.* 10, 165–183.
- Koivisto, M., and Revonsuo, A. (2007). Electrophysiological correlates of visual consciousness and selective attention. *Neuroreport* 18, 753–756.
- Llinás, R. R., and Pare, D. (1991). Of dreaming and wakefulness. *Neuroscience* 44, 521–535.
- Logothetis, N. K. (1998). Single units and conscious vision. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 353, 1801–1818.
- Lou, H. C., Luber, B., Stanford, A., and Lisanby, S. H. (2010). Self-specific processing in the default network: a single-pulse TMS study. *Exp Brain Res.* 207, 27–38.
- Madl, T., Baars, B. J., and Franklin, S. (2011). The timing of the cognitive cycle. *PLoS ONE* 6:e14803. doi:10.1371/journal.pone.0014803
- Madl, T., and Franklin, S. (2012). “A LIDA-based model of the attentional blink,” in *Proceedings of the 11th International Conference on Cognitive Modeling*, (ICCM 2012).
- Maril, A., Wagner, A. D., and Schacter, D. L. (2001). On the tip of the tongue: an event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron* 31, 653–660.
- Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science* 309, 2228–2232.
- Milner, A. D., and Goodale, M. A. (2008). Two visual systems re-reviewed. *Neuropsychologia* 46, 774–785.
- Modha, D. S., and Singh, R. (2010). Network architecture of the long-distance pathways in the macaque brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13485–13490.
- Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., et al. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: a unified account based on multiple trace theory. *J. Anat.* 207, 35–66.
- Nadel, L., Hupbach, A., Gomez, R., and Newman-Smith, K. (2012). Memory formation, consolidation and transformation. *Neurosci. Biobehav. Rev.* 36, 1640–1645.
- Nadel, L., Samsonovich, A., Ryan, L., and Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* 10, 352–368.
- Newell, A. (1990). *Unified Theories of Cognition*, Cambridge: Harvard University Press.
- Panagiotaropoulos, T. I., Deco, G., Kapoor, V., and Logothetis, N. K. (2012). Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron* 74, 924–935.
- Pannu, J. K., and Kaszniak, A. W. (2005). Metamemory experiments in neurological populations: a review. *Neuropsychol. Rev.* 15, 105–130.
- Rosanova, M., Gosseries, O., Casarotto, S., Boly, M., Casali, A. G., Bruno, M. A., et al. (2012). Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain* 135, 1308–1320.
- Ryan, L., Lin, C. Y., Ketcham, K., and Nadel, L. (2010). The role of medial temporal lobe in retrieving spatial and nonspatial relations from episodic and semantic memory. *Hippocampus* 20, 11–18.
- Ryan, L., Nadel, L., Keil, K., Putnam, K., Schnyer, D., Trouard, T., et al. (2001). Hippocampal complex and retrieval of recent and very remote autobiographical memories: evidence from functional magnetic resonance imaging in neurologically intact people. *Hippocampus* 11, 707–714.
- Schneider, W. (2009). “Automaticity and consciousness,” in *Encyclopedia of Consciousness*, 1st Edn, ed. W. Banks (Amsterdam: Academic Press), 83–92.
- Shanahan, M. (2010). *Embodiment and the Inner Life*. Oxford: Oxford University Press.
- Shanahan, M. (2012). The brain’s connective core and its role in animal cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2704–2714.
- Shanahan, M., and Baars, B. (2005). Applying global workspace theory to the frame problem. *Cognition* 98, 157–176.
- Sherman, S., and Guillory, R. (2011). Distinct functions for direct and transthalamic corticocortical connections. *J. Neurophysiol.* 106, 1068–1077.
- Shimamura, A. P. (2010). Hierarchical relational binding in the medial temporal lobe: the strong get stronger. *Hippocampus* 20, 1206–1216.
- Snaider, J., McCall, R., and Franklin, S. (2011). The LIDA Framework as a General Tool for AGI. *Fourth Conference on Artificial General Intelligence*. Mountain View, CA: Springer Lecture Notes in Artificial Intelligence.
- Sporns, O., and Honey, C. (2006). Small worlds inside big brains. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19219–19220.
- Steriade, M. (2006). Grouping of brain rhythms in corticothalamic systems. *Neuroscience* 137, 1087–1106.
- Valderrama, M., Crépon, B., Botella-Soler, V., Martinerie, J., Hasboun, D., Alvarado-Rojas, C., et al. (2012). Human gamma oscillations during slow wave sleep. *PLoS ONE* 7:e33477. doi:10.1371/journal.pone.0033477
- Wang, X. J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455–463.
- Zeki, S. (2001). Localization and globalization in conscious vision. *Annu. Rev. Neurosci.* 24, 57–86.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 November 2012; accepted: 02 April 2013; published online: 28 May 2013.

*Citation: Baars BJ, Franklin S and Ramsoy TZ (2013) Global workspace dynamics: cortical “binding and propagation” enables conscious contents. *Front. Psychol.* 4:200. doi: 10.3389/fpsyg.2013.00200*

This article was submitted to Frontiers in Consciousness Research, a specialty of Frontiers in Psychology.

Copyright © 2013 Baars, Franklin and Ramsoy. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

BASIC EVIDENCE

1. Raw EEG signature: conscious states show irregular, low-amplitude, and fast EEG activity 12–200 Hz, using both intracranial and surface measures. Unconscious states tend to show regular, high-amplitude, and slow voltages at less than 4 Hz.
2. Cortex and thalamus: consciousness depends on the thalamocortical complex, turned on and off by brainstem neuromodulation. Cortical regions are bidirectionally linked to corresponding layered arrays in thalamus.
3. Widespread brain activity: conscious contents are associated with widespread brain activation related to the content. Unconscious stimulation evokes only local cortical activity. Conscious scenes also involve wide effects outside the focus of current conscious contents, as indicated by implicit learning, episodic memory, biofeedback training of autonomic and motor function, and the like.
4. Wide range of contents: consciousness has an extraordinary range of different contents – perception in the various senses, endogenous imagery, emotional feelings, inner speech, concepts, action-related ideas and “fringe” experiences such as feelings of familiarity.
5. Informativeness: consciousness may fade when signals become redundant; a loss of information may also lead to a loss of conscious access. Studies of attentional selection also show a strong preference for more informative conscious stimuli.
6. The rapidly adaptive and fleeting nature of conscious events: immediate experience of the sensory past may last a few seconds, and our fleeting present is perhaps 10–30 s, the duration of working memory when arbitrary material can be accessed at will. In contrast, vast bodies of unconscious knowledge reside in long-term memory.
7. Internal consistency: consciousness is marked by a consistency constraint. For example, while multiple meanings of most words are active for a brief time after presentation, only one becomes conscious at any moment. In general, of two mutually inconsistent stimuli presented simultaneously, only one can become conscious.
8. Limited capacity and seriality: the capacity of consciousness at any given moment seems limited to one consistent scene (see above). The flow of such scenes is serial, in contrast with the massive parallelism of the brain as observed directly.
9. Sensory binding: the sensory brain is functionally segregated such that different cortical areas are specialized to respond to different features such as shape, color, or object motion. One classic question is how these functionally segregated regions coordinate their activities in order to generate the gestalts of ordinary conscious perception.
10. Self-attribution: conscious experiences are always attributed to an experiencing self, the “observing self” as James called it. Self-functions appear to be associated with midline regions of cortex in humans.
11. Accurate reportability: conscious contents are reportable by a wide range of voluntary responses, often with very high accuracy. The standard operational index of consciousness is based on accurate reportability.
12. Subjectivity and privacy: consciousness is marked by the existence of a private flow of events available only to the experiencing subject, though much of it is available for public report.
13. Focus-fringe structure: while consciousness tends to be thought of as a focal, clearly articulated set of contents, “fringe conscious” events, like feelings of familiarity, the TOT experience, etc., are also important.
14. Facilitation of learning: there is very little evidence for long-term learning of unconscious input. In contrast, the evidence of learning of conscious episodes is overwhelming. Even implicit learning requires conscious attention to the stimuli from which implicit regularities are (unconsciously) inferred.
15. Stability of contents: conscious contents are impressively stable, given the variability of input that is dealt with. Even abstract contents such as beliefs, concepts, and the motivational self are remarkably stable over years.
16. Allocentricity: neural representations of external objects make use of diverse frames of reference. Conscious scenes, generally speaking, have allocentric character, though they are shaped by egocentric and other unconscious frameworks.
17. Conscious knowing and decision making: consciousness is obviously useful for knowing the world around us, as well as for knowing certain of our own internal processes. Conscious intentionality may be particularly well suited for voluntary decision making.
18. Orientation to time, place, self, and other.
19. Simple working memory.
20. Intermediate recall.
21. Attention, mental effort, and calculation.
22. Simple language, such as naming common objects.
23. Basic visuomotor skills.

Reflections and Reviews

Losing Consciousness: Automatic Influences on Consumer Judgment, Behavior, and Motivation

JOHN A. BARGH*

Consumer research has largely missed out on two key developments in social cognition research: the growing evidence that much of social judgment and behavior occur without conscious awareness or intent and the substantial moderating influence of social- and self-related goal pursuits on basic cognitive and reasoning processes. This evidence is described and its implications are drawn for nonconscious—including subliminal— influences on consumer behavior. The consumer research domain appears ideal for the necessary next wave of this research: the assessment of how much of a role nonconscious influences play in real life in decisions and behavior that are of real consequence to the individual.

To what extent are people aware of and in control of the influences and reasons for their purchasing and consumption behavior? Although in the past decade of consumer research there has been increasing attention to the possibility that there may be automatic or nonconscious influences on choices and behavior, the field still appears dominated by purely cognitive approaches, in which decisions and actions are made deliberately. Not only does the role of nonconscious processes seem underplayed, relative to contemporary social psychological models, so too do directive motivational influences on reasoning and behavior. My own goal in this article is to describe briefly the recent developments in automatic and nonconscious research in social cognition, consider their relevance to consumer behavior, and then consider their implications for future directions in consumer research.

I am not the first to point to the nonconscious nature of

much of human behavior (see, e.g., Alba 2000; Loewenstein 1996). However, much of the previous discussion within consumer research has treated nonconscious influences as hedonic impulses, which, if they reach expression in behavior, reflect a failure of volitional control or a weakness of will (Alba 2000, p. 3; Baumeister 2002). Instead, I seek to expand the consideration of nonconscious motivations beyond hedonic impulses and physiological need states (such as in addiction) to the operation of any kind of goal or motivation a person can have consciously, such as self-protective motivation, performance- or achievement-related motivation, and interpersonal goals.

According to the recent major surveys of consumer research (Cohen and Chakravarti 1990; Jacoby, Johar, and Morrin 1998; Simonson et al. 2001), the major emphasis is on purchase decisions, with a dominant cognitive approach to understanding how they are made. This cognitive orientation takes two main forms, social cognition and behavioral decision theory. The most influential social cognition models are the elaboration likelihood model (ELM; Petty, Cacioppo, and Schumann 1983) and the heuristic-systematic model (HSM; Chaiken 1980); the predominant paradigm of decision research is an information-processing model. All of these major approaches posit consciously made, deliberate choices and decisions. The social cognition models are

*John A. Bargh is professor and director of the Graduate Program in Social Psychology at New York University, 6 Washington Place, Seventh Floor, New York, NY 10003 (john.bargh@nyu.edu). This essay was prepared while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences and supported as well by a Guggenheim Fellowship and by grant MH-60767 from the National Institute of Mental Health. He thanks Rashmi Adaval, Gavan Fitzsimons, and David Mick for their advice and comments on an earlier version.

mainly concerned with the conditions under which people do, versus do not, engage in careful, effortful processing of the information contained in persuasive messages. In this research, even when the experimental participant is not processing effortfully but instead is relying on simplifying shortcuts, she is still focusing her attention on the message, intentionally processing it, and then reporting a consciously formed attitude or opinion.

Since 1980, however, there have been two important developments in social cognition research that would appear to be highly relevant to consumer research. These are the substantial role played by nonconscious processes (and the minimal role played by deliberate, effortful processes) in psychological and behavioral phenomena and the central and modifying role of needs and goal pursuits. Because of the continued reliance of consumer research on the 1980-era models, these new developments have not yet had much impact. However, to the extent that consumers are behaving without conscious awareness and guidance, models that assume the consumer's deliberate and effortful scrutiny of the choice or behavior will likely miss much of the character and flavor of consumer behavior *in situ*.

CONSCIOUSNESS DETHRONED

The early social-cognitive models were based mainly on research in controlled, quiet, distraction-free settings in which the participant gave his or her full attention to the particular stimuli presented, had plenty of time to consider his or her response, and tried reasonably hard to follow the instructions given explicitly by the experimenter. But soon, researchers began to ask to what extent these models applied to the real, complex, noisy, and busy world outside of the laboratory and to what extent people spontaneously pursued the particular goals given to them by psychology experimenters. As it turned out, the main feature that dropped out of these models as they were studied under more naturalistic conditions was the role played by deliberate conscious choice processes (see Bargh and Chartrand 1999 for a review). For example, Fazio et al. (1986) showed that evaluations of the objects and events in one's environment were made in an immediate, automatic manner, upon the mere presence or occurrence of that object or event. Causal attributions were found to be made substantially through automatic and nonconscious means (e.g., Gilbert 1989), and behavior in social interactions was found to be influenced automatically by social-perceptual mechanisms (Bargh, Chen, and Burrows 1996).

The most recent developments have been in the area of social behavior and goal pursuit. What is most striking about these varied demonstrations of nonconscious influences is that the effects are obtained by the mere, passive activation of the relevant mental concepts, such as intelligent, polite, power, cooperation, and achievement. Most research has produced the activation of these concepts through priming manipulations that typically involve exposure to the concept and close synonyms in the context of an allegedly unrelated prior experiment.

For example, if you present a subject with adjectives related to politeness, in the course of an ostensible language test in which she constructs grammatical sentences out of series of words presented in a scrambled order, and then give her a chance to behave in a polite manner (e.g., waiting patiently for the experimenter to end a conversation with another person), she will exhibit greater politeness (i.e., wait longer before interrupting) than will participants in a control condition (Bargh et al. 1996, experiment 1).

Motivations and goal pursuits can be activated and put into operation in the same way (Bargh et al. 2001). For example, priming the concept of achievement causes participants to work harder and score higher on a verbal task, while priming the concept of cooperation causes them to return more of a common resource (fish in a smallish lake) in order to maintain and continue that resource for all. Not only are people unaware that these goals have been activated but they also are unaware of their operation, even though they are behaving in ways to attain that very goal. Moreover, nonconscious goal pursuits display all of the same features as does (volitional) conscious goal pursuit, such as flexibility, persistence, and effects of success and failure on mood (Chartrand and Bargh 2002).

One might object that these findings are only the results of clever tricks in contrived experimental settings and are not typical of normal, "real world" functioning. Indeed, Clore and Ketelaar (1997) have suggested that these demonstrations of nonconscious influences are analogous to hot-wiring a car: "Automobiles, for example, are not designed to be started with a screw driver and wire clippers, but we all know that the design of a car allows such hot-wiring to happen" (p. 116). To this reasonable point two rejoinders can be made. First, the evidence in support of nonconscious goal pursuit is obtained under conditions that as much as possible mimic those in the real, nonlaboratory world. Ideally, the dependent measure is taken when the participant believes he or she is entirely outside of an experimental situation—when arriving, when between different studies, or when leaving the lab. In one study, for example (Bargh et al. 1996, experiment 2), participants were first primed, or were not primed, with stimuli related to the elderly stereotype, and the dependent measure was how quickly they walked down the hall when leaving the experiment. Other studies of nonconscious goal pursuit employed natural settings that were expected to automatically activate that goal, such as priming power-related goals by having the participant sit in the professor's versus the guest chair in the professor's office (Chen, Lee-Chai, and Bargh 2001).

Second, the original priming studies from the 1960s and 1970s were not laboratory studies but were carried out in the field. These concerned the impact of the presence of aggressive cues such as weapons on subsequent aggressive behavior (Turner, Layton, and Simons 1975), the impact of exposure to television violence on aggressiveness in children (Belson 1978), and the impact of witnessing helpful acts on subsequent likelihood to help a person in need, such as a stranded motorist (Bryan and Test 1967). Those studies

showed consistently strong influences of contextual primes on behavior in real world settings; indeed, the model of aggressive cues developed in the laboratory holds even more strongly when it is tested in the field (Bushman and Anderson 1998).

Still, much more needs to be done to test the extent of influence that the recently discovered forms of nonconscious processes have in normal, everyday functioning. The realm of consumer research would seem to be the ideal playing field on which to establish whether the new models of automatic goal pursuit and automatic evaluation processes do, indeed, apply in the real world, for consumer research involves the study of circumstances in which the person is motivated and involved, where his or her money is on the line, and where the outcome of choices matters to his or her health or happiness.

EXPANDING THE STUDY OF CONSUMER MOTIVATIONS

In order to bring the study of nonconscious motivation into the consumer realm, the range of consumer motivations that are considered needs to be expanded. Perhaps as a legacy of the influential ELM and HSM models of the early 1980s, the motivation to engage (or not to engage) in effortful consideration of a decision or judgment (based on economic calculations of personal cost and benefit for doing so) seems to be the dominant kind of goal or motive studied in contemporary consumer research (for an important recent exception, see Ratneshwar, Mick, and Huffman 2000). But in the real world, people have many other goals and needs. Aside from (occasionally) being motivated to effortfully process product-relevant information or advertising content or to deliberately make choices about which products to buy, people have things they need to get done and pressing concerns on their minds. We want others to like us and to want to be with us; we want to perform well and to achieve success, to ensure our family's safety and security, and to present ourselves to others in a positive, attractive light. Social cognition research over the past 20 years has emphasized the moderating role of these varied motivations for cognitive processes such as attention and judgment (e.g., Tesser, Martin, and Cornell 1996).

The importance of including these various motivations in the research mix is that the particular goal in place changes everything—the focus of attention and the evaluation of objects and events, as well as memory for events (e.g., Bruner 1957; Lazarus 1991). For instance, the currently operating goal drives appraisal and evaluation of objects and events in the current environment (Ferguson and Bargh, forthcoming). To the extent that these help to satisfy the goal, they are positively evaluated and approached; to the extent that they thwart or interfere with goal pursuit, they are negatively evaluated and avoided. Activate intimacy needs or goals, and then products that can help one become more attractive, such as cosmetics or grooming products, should be more positively evaluated than otherwise; activate

health-related goals, and the person should evaluate groceries in terms of his or her health values and their implications; activate gratification or hedonistic goals, and the food's tastiness will dominate the evaluations (see Ramamathan and Menon 2001). Competition goal operation should cause one to evaluate status-oriented products more positively, while egalitarian or responsibility-related goals would cause a more negative evaluation of those same products, and so on.

My point is that all of these goals can be activated, and then operate, all outside of awareness. If they are operating nonconsciously, the person will not be aware of their influence on his evaluations and behavior. The above research shows that the technology and knowledge now exist to activate these goals without the person's awareness.

NONCONSCIOUS INFLUENCE: SPECTER AND REALITY

There are two ways to deliver nonconscious primes: either subliminally, in which case the primes themselves are not accessible to the person's awareness, or supraliminally, in which case the person is aware of the primes but not of their potential influence. Both forms have been shown to be successful in influencing judgments, motivations, and behaviors in social cognition research (Bargh 1992).

Subliminal advertising and subliminal influence attempts more generally have a controversial and checkered past. Indeed, one reason why consumer research seemed to shy away from the study of motivational influences over the past 40 years is the legacy of Vance Packard's 1957 book *The Hidden Persuaders*. Packard proclaimed that market researchers of that time were able to determine people's unconscious motives and that the consumer was powerless to resist these techniques. As evidence, he trumpeted claims of powerful subliminal advertising effects in movie theaters. The book, published in an era of prisoner of war brainwashing attempts and cold war paranoia, was a sensation and gave the scientific study of consumer motivations an unsavory public image.

However, the early reports of subliminal ads in movie theaters turned out to be a hoax, and then the first reviews of the effectiveness of subliminal advertising showed weak effects at best (Moore 1982). Next, experimental tests of the effectiveness of commercially available subliminal self-help tapes found them to be no more effective than placebo tapes (Greenwald et al. 1991).

Today, most people remain concerned about the possibility of being influenced by subliminal messages (Wilson and Brekke 1994), and perhaps now, finally, they should be. Contemporary researchers are consistently obtaining subliminal effects on consumption and product evaluation. What has changed?

The main reason for the recent success is that researchers are taking the consumer's (experimental participant's) cur-

rent goals and needs into account.¹ Lewin's influential field theory (1951) held that one could not induce in people goals they do not already have themselves, but you could influence them by activating or manipulating the goals that they already possessed. The most recent work on subliminal influence exploits this principle by matching the subliminal stimulus with the subject's current goal or need state; it also makes use of known effective primes. Thus, Berridge and Winkielman (forthcoming) subliminally presented subjects with a happy, a neutral, or an angry face. Those who had been shown the happy face subsequently evaluated a fruit-flavored drink more favorably and also drank substantially more of it than did neutral-primed participants. Those who had been shown the angry face drank least of all. Most important, these effects held only for those participants who were thirsty (having been instructed not to drink anything for hours before the experiment); the evaluations and drinking behavior of nonthirsty participants were unaffected by the same subliminal primes.

Strahan, Spencer, and Zanna (forthcoming) subliminally primed thirstiness and caused thirsty, but not nonthirsty, participants to drink more of a purportedly thirst-quenching beverage ("SuperQuencher") than of a purportedly energy-giving beverage ("PowerPro"). In another study, they primed sadness and thereby caused participants to prefer listening to a CD purported to put them in a good mood rather than to a CD of music described as strong and powerful. Again, in both studies non-goal-primed participants did not show these effects. For the subliminal effects to occur required the match between the needs and goals of the participant and the needs that the product was alleged to satisfy (see also Dijksterhuis et al., forthcoming).

With subliminal primes, the individual has no chance of controlling the influence; as they used to say of Bob Feller's fastball, you can't hit what you can't see. But most stimuli in real life as well as in advertising are in one's plain view. Supraliminal influence attempts, including goal activation, can be as effective—if not more effective—than subliminal priming. What is critical is that people not be aware of how the primes might affect them. Given our general overconfidence in our ability to be aware of the important influences on our judgments and behavior, as well as in our ability to control any unwanted influence, this condition is met much of the time. After all, almost all external, environmental influences on our behavior involve stimuli and messages that are in plain view, yet either we do not realize that the influence is taking place (Wilson and Brekke 1994) or we are overconfident as to our ability to control any such influences (Bargh 1999b). As an example, we routinely bemoan negative or "dirty" political campaign advertising and insist that such ads do not affect our own vote, yet the reason such ads do not go away, and even increase in frequency each election season, is that they are, in fact, quite effective.

¹The other reason is that priming researchers are using techniques that work, such as multiple presentations of the subliminal prime instead of just one and using single words instead of entire sentences that need to be parsed (see Dijksterhuis, Aarts, and Smith, forthcoming).

There are, however, situations and contexts in which we are aware of attempts to influence us and even of how that attempt might operate. Millions of people watch the Super Bowl football telecast in part to see the new (and very expensive) ads, fully aware while watching them that they are trying very hard to influence us in some way. We know full well that repeated ads asking, "What is mLife?" without ever telling us are trying to incite us to go to the displayed internet address to find out (and so we can resist this impulse). These are the very cases in which subliminal influence attempts will likely prove to be more successful than supraliminal (conscious) ones.

THE USE AND ABUSE OF NONCONSCIOUSNESS

It would be naive to think that recent advances in knowledge of nonconscious processes will never be exploited to serve a company's or a government's purposes against the interests of consumers or citizens. Therefore, within the field of consumer research, there are choices to be made concerning how to approach the study of nonconscious influences. Here the question neatly becomes, Who is the intended consumer of that research? Who is the master being served?

Nearly 40 years ago, Robert Perloff (1964) wrote that in the first half-century of industrial psychology's existence, the overwhelming amount of research treated the consumer "as an individual whose attention and purchasing behavior are coveted to serve ends . . . determined by advertising and the mass media" (p. 33). According to the recent reviews of the field, not much has changed, and most research still is devoted to influences on purchase decisions (Simonson et al. 2001, p. 255). Perloff concluded that the benefits of consumer research for the individual in society would be greatly multiplied if researchers directly and explicitly targeted the consumer's motives and needs.

In Perloff's day, the idea of "hidden persuaders" was indeed a matter of bluff and, to some extent, hysteria. But no longer. If researchers had a responsibility in 1964 to serve the consumer, how much more of a responsibility is there today, when methods to thwart or bypass the consumer's defenses against influence are becoming ever more powerful, and yet he remains as ignorant of these influences and as overconfident of his control as in the past?

There is a way that contemporary consumer researchers can live up to this responsibility. In social cognition research, the 1980s saw growing evidence of the nonconsciousness of stereotyping and prejudice—of these being automatic, unintended, and possibly even uncontrollable influences on judgment and behavior toward the stereotyped group. The response of the field was a massive research effort, that continues today, into ways in which people could overcome these influences and regain control (see Bargh 1999a). Perhaps consumer research should begin to balance studies of how to influence the consumer's choices and behavior with studies of how she can defend against and control such

unwanted influences. Researching both sides of the issue of nonconscious influence would also provide the strongest and most relevant evidence to date on the basic and important research question of how powerful and typical nonconscious influences are in daily life, because consumer research is the study of choices and behaviors that really matter to the individual.

[David Glen Mick served as editor for this article.]

REFERENCES

- Alba, Joseph W. (2000), "Dimensions of Consumer Expertise . . . or Lack Thereof," *Advances in Consumer Research*, 27, 1-9.
- Bargh, John A. (1992), "Why Subliminality Does Not Matter to Social Psychology: Awareness of the Stimulus versus Awareness of Its Influence," in *Perception without Awareness*, ed. Robert F. Bornstein and Thane S. Pittman, New York: Guilford, 236-255.
- (1999a), "The Cognitive Monster: The Case against Controllability of Automatic Stereotype Effects," in *Dual Process Theories in Social Psychology*, ed. Shelly Chaiken and Yaacov Trope, New York: Guilford, 361-382.
- (1999b), "The Most Powerful Manipulative Messages Are Hiding in Plain Sight," *Chronicle of Higher Education* (January 29), B6.
- Bargh, John A., and Tanya L. Chartrand (1999), "The Unbearable Automaticity of Being," *American Psychologist*, 54 (July), 462-479.
- Bargh, John A., Mark Chen, and Lara Burrows (1996), "Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Priming on Action," *Journal of Personality and Social Psychology*, 71 (September), 230-244.
- Bargh, John A., Peter M. Gollwitzer, Annette Lee-Chai, Kim Barn-dollar, and Roman Troetschel (2001), "The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals," *Journal of Personality and Social Psychology*, 81 (December), 1014-1027.
- Baumeister, Roy F. (2002), "Yielding to Temptation: Self-Control Failure, Impulsive Purchasing, and Consumer Behavior," *Journal of Consumer Research*, 28 (March), 670-676.
- Belson, William A. (1978), *Television Violence and the Adolescent Boy*, Westmead: Saxon House.
- Berridge, Kent C., and Piotr Winkielman (forthcoming), "What Is an Unconscious Emotion? (The Case for Unconscious 'Liking')," *Cognition and Emotion*.
- Bruner, Jerome S. (1957), "On Perceptual Readiness," *Psychological Review*, 64 (January), 123-152.
- Bryan, James H., and Test, Mary A. (1967), "Models and Helping: Naturalistic Studies in Aiding Behavior," *Journal of Personality and Social Psychology*, 6 (August), 400-407.
- Bushman, Brad J., and Craig A. Anderson (1998), "Methodology in the Study of Aggression: Integrating Experimental and Nonexperimental Findings," in *Human Aggression: Theories, Research, and Implications for Policy*, ed. Russell Geen and Edward Donnerstein, San Diego, CA: Academic Press, 23-48.
- Chaiken, Shelly (1980), "Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion," *Journal of Personality and Social Psychology*, 39 (November), 752-766.
- Chartrand, Tanya L., and John A. Bargh (2002), "Nonconscious Motivations: Their Activation, Operation, and Consequences," in *Self and Motivation: Emerging Psychological Perspectives*, ed. Abraham Tesser, Diederik A. Stapel, and Joanne V. Wood, Washington, DC: American Psychological Association, 13-41.
- Chen, Serena, Annette Y. Lee-Chai, and John A. Bargh (2001), "Relationship Orientation as a Moderator of the Effects of Social Power," *Journal of Personality and Social Psychology*, 80 (February), 173-187.
- Clore, Gerald, and Timothy Ketelaar (1997), "Minding Our Emotions: On the Role of Automatic, Unconscious Affect," in *Advances in Social Cognition*, Vol. 10, ed. Robert S. Wyer, Jr., Mahwah, NJ: Erlbaum, 105-120.
- Cohen, Joel B., and Dipankar Chakravarti (1990), "Consumer Psychology," *Annual Review of Psychology*, 41, 243-288.
- Dijksterhuis, Ap, Henk Aarts, and Pamela K. Smith (forthcoming), "The Power of the Subliminal: On Subliminal Persuasion and Other Potential Applications," in *The New Unconscious*, ed. Ran Hassin, James Uleman, and John Bargh, New York: Oxford University Press.
- Fazio, Russell H., David M. Sanbonmatsu, Mary C. Powell, and Frank R. Kardes (1986), "On the Automatic Activation of Attitudes," *Journal of Personality and Social Psychology* 50 (February), 229-238.
- Ferguson, Melissa J., and John A. Bargh (forthcoming), "Sensitivity and Flexibility: Exploring the Knowledge Function of Automatic Attitudes," in *The Wisdom of Feelings: Processes Underlying Emotional Intelligence*, ed. Lisa F. Barrett and Peter Salovey, New York: Guilford.
- Gilbert, Daniel T. (1989), "Thinking Lightly about Others: Automatic Components of the Social Inference Process," in *Unintended Thought*, ed. James S. Uleman and John A. Bargh, New York: Guilford, 189-211.
- Greenwald, Anthony G., Eric R. Spangenberg, Anthony R. Pratkanis, and Jay Eskenazi (1991), "Double-Blind Tests of Subliminal Self-Help Audiotapes," *Psychological Science*, 2 (March), 119-122.
- Jacoby, Jack, Gita V. Johar, and Maureen Morrin (1998), "Consumer Behavior: A Quadrennium," *Annual Review of Psychology*, 49, 319-344.
- Lazarus, Richard S. (1991), *Emotion and Adaptation*, New York: Oxford University Press.
- Lewin, Kurt (1951), *Field Theory in Social Science*, New York: Harper & Row.
- Loewenstein, George F. (1996), "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, 65 (March), 272-292.
- Moore, Timothy E. (1982), "Subliminal Advertising: What You See Is What You Get," *Journal of Marketing*, 46, 38-47.
- Packard, Vance (1957), *The Hidden Persuaders*, New York: David McKay.
- Perloff, Robert (1964), "Potential Contributions of the Consumer-Oriented Psychologist," *Business and Society*, 4 (Spring), 28-34.
- Petty, Richard E., John T. Cacioppo, and David Schumann (1983), "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement," *Journal of Consumer Research*, 10 (September), 135-146.
- Ramanathan, Suresh, and Geeta Menon (2001), "Don't Know Why but I Had This Craving: Goal Dependent Automaticity in Impulsive Decisions," Working Paper No. 11-01, Stern School of Business, New York University, NY 10012.

- Ratneshwar, S., David G. Mick, and Cynthia Huffman, eds. (2000), *The Why of Consumption*, New York: Routledge.
- Simonson, Itamar, Ziv Carmon, Ravi Dhar, Aimee Drolet, and Stephen M. Nowlis (2001), "Consumer Research: In Search of Identity," *Annual Review of Psychology* 52, 249–275.
- Strahan, Erin J., Steven J. Spencer, and Mark P. Zanna (forthcoming), "Subliminal Priming and Persuasion: Striking While the Iron Is Hot," *Journal of Experimental Social Psychology*.
- Tesser, Abraham, Leonard L. Martin, and David P. Cornell (1996), "On the Substitutability of Self-Protective Mechanisms," in *The Psychology of Action*, ed. Peter M. Gollwitzer and John A. Bargh, New York: Guilford, 48–68.
- Turner, Charles W., John F. Layton, and Lynn S. Simons (1975), "Naturalistic Studies of Aggressive Behavior: Aggressive Stimuli, Victim Visibility, and Horn Honking," *Journal of Personality and Social Psychology*, 31 (June), 1098–1107.
- Wilson, Timothy D., and Nancy Brekke (1994), "Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations," *Psychological Bulletin*, 116 (July), 117–142.

Nonconscious Goals and Consumer Choice

TANYA L. CHARTRAND
JOEL HUBER
BABA SHIV
ROBIN J. TANNER*

This work examines the process through which thrift versus prestige goals can nonconsciously affect decisions in a choice task. Drawing upon research on nonconscious goal pursuit, we present a theoretical framework detailing how consumer choices are affected by incidentally activated goals. We show that such primed goals have motivational properties consistent with goal pursuit but inconsistent with mere cognitive activation; the effects are greater with a longer time interval between the priming task and the choice and are less pronounced when the primed goal is satiated in a real, as opposed to a hypothetical, intervening choice task. Additionally, we show that subliminally evoked retail brand names can serve as the cues that activate purchasing goals.

Consider the following scenario: You are at a local shopping center to purchase crew socks from a sports store. On the way to the sports store, you pass by either a Nordstrom, where you are incidentally exposed to images of prestige (e.g., luxury, extravagance), or a Wal-Mart, where you are incidentally exposed to images of thrift (e.g., "everyday low prices"). The sports store carries two brands of crew socks, one more expensive and prestigious (e.g., Nike), and the other a better value (e.g., Hanes). The question is, does your incidental exposure to Nordstrom (Wal-Mart) activate a nonconscious prestige-oriented (thrift-oriented) goal that increases the likelihood that you will choose Nike (Hanes)? The goal of the present research is to explore this intriguing possibility. Drawing upon research on nonconscious goal pursuit (e.g., Bargh et al. 2001; Chartrand and Bargh 1996, 2002; Chartrand, Dalton, and Cheng, forthcoming), we examine

whether incidental exposure to cues can activate different shopping goals and, in turn, influence subsequent decisions on unrelated choice tasks in a nonconscious manner.

It is now widely accepted from research spanning three decades that consumer behavior is largely goal-directed. Goals are credited with being a key motivational construct guiding consumer decision making (Bettman 1979; Bettman, Luce, and Payne 1998; Fishbach and Dhar 2005; Higgins 2002; Kivetz, Urmansky, and Zheng 2006; Shiv and Huber 2000; Soman and Cheema 2004). Despite the preponderance of research on goals in marketing, several scholars have noted that little research on goals has explicitly examined the factors that determine how consumer goals actually come to be selected and pursued (Bagozzi and Dholakia 1999; Huffman and Houston 1993). In a rare exception, Bagozzi and Dholakia (1999) outline a conceptual framework for thinking about how goal setting and goal pursuit influence consumer behavior. However, this framework deals primarily with the conscious aspects of goal pursuit and leaves aside the possibility of nonconscious factors influencing goal pursuit.

In recent years, a growing body of research has documented the possibility of nonconscious goal pursuit, challenging the traditional view that consumers are fully cognizant of the goals underlying their decisions (e.g., Bargh 2002; Chartrand et al., forthcoming). Specifically, research on nonconscious goal pursuit suggests that goals can be activated by situational cues and can influence behavior outside of awareness until the desired outcome has been attained (Chartrand and Bargh 1996). This form of goal pursuit operates without need for conscious intervention but has behavioral consequences that parallel those arising from consciously set goals.

*Tanya L. Chartrand is associate professor of marketing and psychology at the Fuqua School of Business, Duke University, Durham, NC 27708-0120 (tlc10@duke.edu). Joel Huber is Allan D. Schwartz Professor of Marketing at the Fuqua School of Business, Duke University, Durham, NC 27708-0120 (joel.huber@duke.edu). Baba Shiv is professor of marketing at the Stanford Graduate School of Business, Stanford, CA 94305 (shiv_baba@gsb.stanford.edu). Robin J. Tanner is a PhD candidate in marketing at the Fuqua School of Business, Duke University, Durham, NC 27708 (Rob.Tanner@duke.edu); after July 2008, he can be reached at the Wisconsin School of Business, 975 University Avenue, Madison, WI, 53706 (rtanner@bus.wisc.edu). The authors contributed equally to this research and are listed in alphabetical order. The authors thank Himanshu Mishra and Justin Petersen for their help in administering the studies and coding the debriefings and thank the editor, the associate editor, and the three reviewers for their feedback at various stages of this project.

John Deighton served as editor and Stephen M. Nowlis served as associate editor for this article.

Electronically published April 29, 2008

As stated above, a primary objective of the research presented here is to examine the nonconscious activation and pursuit of consumer goals related to prestige versus thrift. These are fairly abstract-level goals that are likely to be associated with various subgoals (e.g., saving money) and plans of action (e.g., buying sale items). The importance of these goals may differ from individual to individual (some consumers are chronically more concerned with prestige or with thrift) and vary over time or in different purchase contexts. Within that premise, this article seeks to answer the following questions about nonconscious goals. First, do consumers need to be conscious of the activation and pursuit of such goals in order for purchase behaviors to be altered? Or is it possible that these nonconscious goals are sometimes activated and pursued outside of conscious awareness, without consumers realizing the influence of such goals on their purchases? If the latter is the case, would such nonconsciously activated goals share kinship with conscious goals, which are often associated with a relatively complex set of actions? For instance, as time passes without a resolution of the nonconscious goal, would the intensity with which the goal is pursued increase, as it does with conscious goals (see Bargh et al. 2001)? Also in parallel with conscious goals, would nonconscious goals diminish in strength once they have been satiated? Answering these questions is important not only to substantiate any goal-based account of priming effects but also to achieve a more complete understanding of such factors, without which we cannot begin to explore the various ways by which consumers, and indeed marketers, can control or change their actions (Chartrand 2005).

The following section provides a theoretical foundation for answering these questions. Subsequent sections then demonstrate, across four studies, that exposure to prestige or thrift cues does indeed engender consumer behavior consistent with a goal pursuit account. In these studies, we demonstrate that these nonconscious goals work similarly irrespective of whether they are supraliminally (studies 1–3) or subliminally (study 4) activated (Bargh and Chartrand 2000). It should be noted that the current work is the first demonstration of the satiation of nonconsciously pursued goals. Additionally, it yields the important finding that although real choices render such satiation, hypothetical choices do not (study 3). Finally, study 4 constitutes the first demonstration that subliminally induced retail brands can themselves invoke consumer goals.

THEORETICAL BACKGROUND

The theoretical framework proceeds in two broad steps. First, we present a detailed conceptualization of nonconscious goal pursuit, leading to a general hypothesis about the effects on choices of priming prestige-related versus thrift-related goals. Then we discuss two factors that are likely to moderate the effects of these goals on choice: the time interval between the goal prime and the choice, and the satiation of the primed goals prior to the choice task. These moderators enable us not only to delineate some prac-

tical boundaries of the effects of nonconsciously held goals on choice behavior but also to provide support for our goal pursuit conceptualization and reduce the viability of alternative explanations of our findings.

Nonconscious Goal Pursuit

Goals are desired end states (Custers and Aarts 2005; Dijksterhuis, Chartrand, and Aarts, forthcoming) that once activated affect behavior in a wide variety of ways. Exactly how goals affect our behavior is a matter of debate in the literature (Van Osselaer et al. 2005). According to a purely cognitive perspective, goals are mentally represented in the same way as are such constructs as schemas, stereotypes, attitudes, and traits (Bargh 1990; Shah and Kruglanski 2003). Once the mental representation of a particular construct is activated, its associated actions will be triggered and performed. For example, being reminded of a goal to save money should trigger the associated goal-driven behavior. However, other theories have posited that positive valence is an important part of the goal representation and plays a role in the motivational processes underlying goal-directed behavior (Custers and Aarts 2005). Specifically, the end state must be linked to affect in order to lead to motivational behavior. Thus, the consumer must have positive feelings toward the desired end state (saving money) in order for a savings-related cue to drive future behavior. For the purposes of the current research, we are agnostic with respect to the affective component of goal pursuit. From our perspective, the important piece is common to both theoretical views: goals are mentally represented, and their activation leads to goal-directed behavior.

This activation can occur in multiple ways. The most intuitive way occurs when individuals consciously decide to pursue a goal or engage in goal-directed behavior. However, Bargh (1990) argued that there are other ways to activate goal constructs in memory. Specifically, environmental cues can be linked to goals in memory. If a certain goal is often pursued in a certain situation, then eventually the two become linked in memory and the presence of the situational feature will automatically activate the goal through a mechanism that shares kinship with stimulus-response conditioning. Once activated, the goal is pursued as if it were consciously chosen, even though the operation of the goal requires no conscious guidance.

According to Bargh (1990), once a goal (at whatever level of abstraction) is activated, the strategies and plans of action associated with that goal should also be automatically activated and should direct subsequent behavior. Supporting this theory, Aarts and Dijksterhuis (2000) found that goal activation led to the nonconscious activation of specific behavioral strategies previously used to attain the goal. These action plans could guide subsequent behavior without the person's awareness or intent.

Chartrand and Bargh (1996) were the first to demonstrate that nonconscious goals activated by environmental cues operate in a fashion similar to those consciously initiated and pursued. Their demonstration built upon the classic find-

ing of Hamilton, Katz, and Leirer (1980)—that participants given an explicit goal to form an impression of a target person subsequently displayed superior memory for the target's behaviors than did participants specifically tasked with memorizing those behaviors. Chartrand and Bargh replaced the explicit instruction used in the original research with a supraliminal priming technique to activate the impression formation goal. In a scrambled-sentence task, participants were exposed to words related either to impression formation (e.g., to evaluate, to judge) or to memorization (e.g., to retain, to absorb). This task unobtrusively primed participants with either an impression formation goal or a memorization goal (see Srull and Wyer 1979). Despite the absence of explicit instruction to form an impression of the target, the study replicated the Hamilton et al. findings: participants previously exposed to impression formation words demonstrated significantly superior recall compared to participants exposed to memorization words.

Since the original demonstration of the parallel functioning of conscious and nonconscious goals, this basic effect has been replicated across a variety of goals, using a variety of both supraliminal and subliminal priming techniques (Aarts and Dijksterhuis 2000; Bargh et al. 2001; Fishbach, Friedman, and Kruglanski 2003; Fitzsimons and Bargh 2003; Shah 2003; Shah and Kruglanski 2003). A critical feature of this line of research is that extensive debriefing of participants consistently reveals a complete absence of any conscious awareness of the nature or effect of the priming manipulations used. As argued by Chartrand (2005), an individual may be unaware of the source of a process, the process itself, or the outcome of the process. In the case of nonconscious goal pursuit, individuals are unaware of the goal activation (source) and goal pursuit (process) but are usually aware of engaging in a particular behavior (outcome). That individuals can be unaware of the source of a goal that becomes active may not be as surprising as the notion that they can also be unaware of pursuing the goal itself. Yet research has shown that goal pursuit leads to the same outcomes regardless of whether the goals are consciously activated or introduced outside of awareness.

Separating Nonconscious Goals from Trait Priming

Although the evidence supporting nonconscious-goal pursuit is compelling, one could offer alternative explanations for the behavioral effects found. Recall that the most common technique for exploring nonconscious goals is to prime words related to the goal construct and then test for effects on behavior. Such goal-related words may, however, also be activating other constructs, such as stereotypes and traits (Bargh 1997). These nonmotivational constructs have also been shown to automatically alter individuals' behavior, leading, for example, to slower walking after exposure to words related to the elderly, without the individuals' conscious intent or awareness (Bargh, Chen, and Burrows 1996; Dijksterhuis and Bargh 2001). We refer to such temporary

increases in association with a particular behavioral construct as a *trait-priming explanation*. Thus, whenever nonconscious effects on behaviors are studied, it is important to distinguish between motivational goal-related processes and cognitive associational processes.

In response to these concerns, researchers have developed techniques designed to tease apart the motivational effects of goal activation from the nonmotivational effects of trait priming. One method relies on the finding that recently acquired goals initially increase in strength over time unless fulfilled (Atkinson and Birch 1970), whereas once activated, nonmotivational constructs decline in strength over time. We refer to this as the *temporal escalation criterion*. Bargh et al. (2001) were the first to use the predictions for temporal escalation to validate the goal-directed nature of prime-influenced behaviors. Participants primed with performance-goal-related stimuli then engaged either in an impression formation or in a word search task, immediately or after a 5-minute delay. The authors predicted that those engaged in impression formation would be more likely to judge the target as achievement oriented in the near term compared with those who experienced a delay. However, because the word search task reflected goal-directed-action tendencies, performance on it should actually improve with the delay. This temporal escalation was exactly what the authors found. While the perceptual effects of the achievement prime (as measured by the extent to which participants judged the target to be achievement oriented) decayed over the delay period, participant performance in the word search task significantly increased with the delay, a result consistent with a goal pursuit account.

A second method used in the current research to distinguish between goal pursuit and associational processes relies on the fact that goals should decrease in strength once the goal has been achieved (Atkinson and Birch 1970). By contrast, associational processes should, if anything, increase as attention is brought to them through achievement. We refer to this distinction as the *goal-satiation criterion*. Thus far, research has not yet assessed goal satiation for nonconsciously held goals. However, substantial evidence exists that the accessibility of goal-related constructs is enhanced while a consciously held goal is active but is inhibited once the goal is fulfilled (Forster, Lieberman, and Higgins 2005; Marsh, Hicks, and Bink 1998; Marsh, Hicks, and Bryan 1999). Because the operation of nonconsciously held goals has been found to be largely indistinguishable from that of consciously held goals, we expect that similar satiation effects would ensue upon the fulfillment of nonconsciously pursued goals.

An issue also arises with respect to what counts as goal satisfaction. Although the initiation of a goal can be evoked by simply cuing the goal (Chartrand and Bargh 1996), it is not clear whether satiation can be as easily evoked. Satiation may require actual satisfaction, particularly if the underlying need has not gone away. For example, attempting to satiate hunger through hypothetical meals will only make the situation worse. We propose that hypothetical satisfaction of

thrift and prestige goals should not work to satiate—rather, that real purchases would be needed to satiate such goals.

Overview of Studies

Four studies explore the effects of priming goals related to prestige versus thrift on hypothetical and real choices. Study 1 demonstrates the basic effect of nonconscious goal activation on choice. Study 2 tests the temporal escalation criterion by manipulating the time delay between goal prime and choice. Study 2 also tests the goal-satiation criterion by having participants make a series of sequential choices. The idea behind this choice task is that if each choice successively satiates the primed goal, then the effects of priming ought to diminish from the first to the last choice. Study 3 continues the exploration of goal satiation effects by examining both real and hypothetical choices. Study 4 demonstrates that naturalistic subliminal cues (in this case, retail brands) can serve to prime the goals.

These studies enable us to make the following claims: First, activating consumer choice goals in a purportedly unrelated priming task alters subsequent choices. Second, the observed effects are in line with the theory of nonconscious goal pursuit. Specifically, the effects occur outside of awareness; they initially increase with a greater time interval between the priming and the choice task, and they decline when satisfied by an intervening real and not hypothetical choice. Finally, rather than being restricted to supraliminal experimenter-generated primes, subliminal exposure to environmental features such as brand names can serve as the primes that activate the goals.

STUDY 1: DEMONSTRATING THE EFFECT OF NONCONSCIOUS CONSUMER GOALS

Study 1 provides an initial test of the ability of nonconsciously activated thrift versus prestige goals to influence subsequent consumer choices.

Design and Procedure

Study 1 randomly assigned 51 participants to either a prestige condition or a thrift condition in a single-factor between-subjects design. Participants initially engaged in a scrambled-sentence task, adapted from Chartrand and Bargh (1996), in which they constructed grammatically correct sentences using four from a list of five scrambled words. For the prestige-primed group, words invoking prestige goals were embedded in the list (e.g., “he prestige what want did”; the solution for which is “what did he want”). For the thrift-primed group, words invoking thrift goals were used (e.g., “he frugal what want did”; the solution for which is the same as in the example presented above). Following a 5-minute filler task, participants made a hypothetical choice between two sock options described as follows: “You notice that it is time to throw away your cotton crew socks and buy new ones. You are considering the following two op-

tions: Nike at \$5.25 a pair and Hanes at \$6 for two pairs.” Participants were subsequently thanked and debriefed.

The debriefing involved a funneled questionnaire protocol (e.g., Bargh and Chartrand 2000; Chartrand and Bargh 1996; Fitzsimons and Shiv 2001) that probed the participants for any suspicions they might have had about the relationship between the original goal-priming task and the experimental task. The participants answered general questions, first, about what they thought the point of the experiment was and, second, about whether they thought that one part of the experiment might have affected another part. If a participant did not indicate any suspicion of a connection, he or she was asked to guess how the original scrambled-sentence task might have been related to the later choice tasks.

Results and Discussion

The results were consistent with our hypothesis. A greater proportion of participants chose the higher-priced Nike socks in the prestige condition than in the thrift condition (48.0% vs. 19.2%; $\chi^2 = 4.75, p < .03$). The funneled debriefing indicated that none of the respondents correctly guessed the general purpose of the study or believed that incidental exposure to words might have altered their choice.

These findings provide initial evidence that nonconscious goals can impact consumer behavior outside of participant awareness. However, study 1 cannot distinguish between a motivational goal pursuit and a trait-priming explanation (Dijksterhuis and Bargh 2001). Study 2 was thus designed to discriminate between these explanations.

STUDY 2: THE IMPACT OF PRIMES INCREASES WITH GREATER INTERVENING TIME

The purpose of study 2 was to replicate the findings of study 1 and provide support for our goal pursuit account. Two changes allowed us to test the motivational nature of the study 1 results. First, we manipulated the time interval between the goal-priming task and the choice task, and, second, we presented participants with a sequence of three different choice tasks, each requiring a trade-off of thrift versus prestige. According to our motivational conceptualization, if the observed effects on choice are indeed related to the activation and pursuit of goals, then increasing the time between the goal prime and the subsequent choices ought to strengthen the activated goals, escalating the effects of goal priming on choice. Further, when presented with a series of goal-relevant decisions, the effect of the prime should diminish as the successive choices begin to satisfy the primed goal.

Design and Procedure

Two hundred and forty-nine participants were randomly assigned to one of the four conditions in a 2 (goal prime: prestige vs. thrift) \times 2 (time delay: high vs. low) between-subjects design. The procedure paralleled that used in study

1 except that participants made choices in two additional product categories (apartments and sound systems). As in study 1, participants first engaged in a scrambled-sentence task containing either prestige- or thrift-related words. They then completed a filler task in which they watched a neutral video of either 3 minutes or 8 minutes in length purportedly being pilot-tested for a different study. (In a separate pilot test, participants drawn from the same population as the main study watched the video for either 3 minutes or 8 minutes, as in study 2, and then responded to a 20-item PANAS mood scale [Watson, Clark, and Tellegen 1988]: the results revealed that mood states were no different across the 3-minute vs. 8-minute video conditions ($p > .20$).

After the time delay filler task, participants chose between two options in each of three different product categories. The first task replicated the crew sock choice presented to participants in study 1. In the second task, participants chose between apartment A, with rent at \$810 per month, an excellent view (cityscape and a river), and a bright and sunny atmosphere, and apartment B, with rent at \$490 per month, a poor view (the back of another building), and a somewhat dark and dreary atmosphere. In the final task, participants chose between a Bose sound system priced at \$1,499, which offered very small speakers and a sleek design that was bound to get "Wows from experts and amateurs alike," and a Toshiba sound system priced at \$669, which offered larger speakers and acoustics that were "Rated favorably by experts in home electronics."

Results

Choice. Table 1 presents the proportion of times the more expensive items were chosen. In line with our predictions, an omnibus logit model across all three product categories with goal prime-, time interval-, and two-product-category-specific dummy variables revealed a significant goal prime by time interval interaction ($\chi^2 = 7.6$, $p < .006$) and a significant main effect of goal prime ($\chi^2 = 55.9$, $p < .0001$). It is noteworthy that there were no significant interactions between the product category dummies

and the independent variables of interest—namely, goal prime and time interval (all p 's $> .20$). Finally, as in the previous study, the funnel debriefing once again revealed no respondents who correctly linked the sentence construction and the choice task.

Effect of Time Delay. As shown in figure 1, across the three product categories the pattern of results in the short time interval conditions confirmed our conceptualization. Specifically, choice of the more expensive options was higher in the prestige goal prime condition (52.7%) than in the thrift goal prime condition (34.4%; $\chi^2 = 15.77$, $p < .0001$). Further, as predicted, choice of the more expensive option when prestige was primed increased as the time interval increased from low (52.7%) to high (63.5%; $\chi^2 = 5.45$, $p < .02$). In contrast, choice of the more expensive option when thrift was primed decreased as the time interval increased from low (34.4%) to high (23.5%; $\chi^2 = 6.46$, $p < .01$).

Stated differently, the difference in choice of the expensive option between the prestige and the thrift goal prime conditions was higher when the time interval was high (63.5% vs. 23.5%) than when it was low (52.7% vs. 34.4%; $\chi^2 = 11.13$, $p < .0001$).

Effect of Satiation. An intriguing result was the lack of the expected interactions between the product dummies and the goal prime (p 's $> .75$), suggesting that the impact of the goal prime did not diminish across the three choices. That is to say, we did not observe the expected satiation of the primed goals, whose effect in fact persisted.

Discussion

The results of study 2 provide further support for the notion that activating prestige versus thrift goals in an unrelated task can give rise to nonconscious goal pursuit. First, across three product categories, choice of the more expensive option was higher when the earlier priming task exposed participants to words related to prestige rather than thrift.

TABLE 1

IMPACT OF PRIME INCREASES CHOICE OF THE PRESTIGE OPTIONS WITH GREATER TIME INTERVAL: STUDY 2

	3-minute interval		8-minute interval		Significant effects	χ^2 values
	Prestige prime (%)	Thrift prime (%)	Prestige prime (%)	Thrift prime (%)		
Crew socks	26.8	12.3	38.6	3.6	GP GP × TI	20.1* 4.2*
Apartments	78.6	63.2	86.0	47.3	GP GP × TI	26.3* 4.0*
Sound systems	53.6	35.1	63.2	23.6	GP GP × TI	21.4* 3.5*
Combined choice	52.7	34.4	63.5	23.5	GP GP × TI	55.9* 7.6*

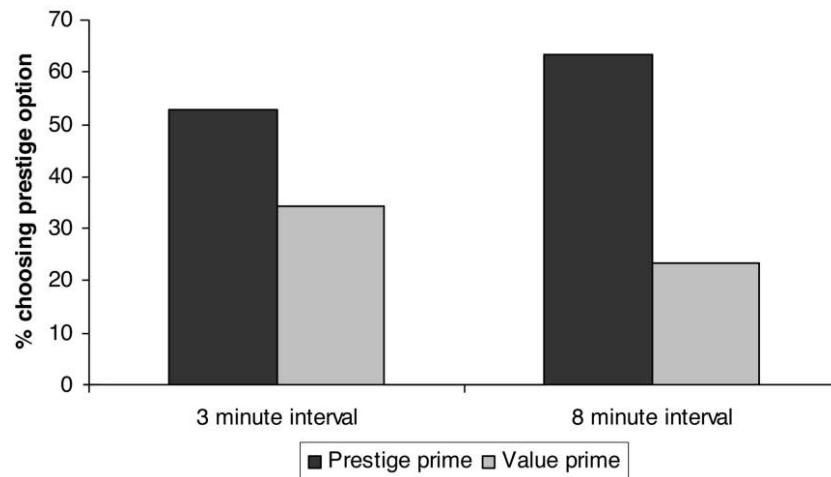
NOTE.—GP = goal prime; TI = time interval.

* $p < .10$.

** $p < .05$.

FIGURE 1

EFFECT OF TIME DELAY ON CHOICE OF THE PRESTIGE OPTION ACROSS ALL THREE CATEGORIES: STUDY 2



Consistent with a goal-priming account but not with a trait-priming account and in line with the temporal escalation criterion, this effect was further augmented with an increase in the time interval between the priming task and the choice tasks. Second, the observed effects seemed to occur outside of awareness inasmuch as extensive questioning of the participants during debriefing revealed no indication of conscious awareness or suspicion of the effect of the priming task on the subsequent choices.

A notable finding is that we did not observe a satiation of goals as the choice task progressed from the first set to the last. On the contrary, the effect of the goal primes persisted consistently across all three choices. A possible explanation for this lack of satiation is that the choices were hypothetical in nature and that real choices are necessary to bring about satiation. Study 3 tests this possibility.

STUDY 3: GOAL SATIATION WITH REAL VERSUS HYPOTHETICAL CHOICES

Study 3 continued the process of replicating the findings of studies 1 and 2. In addition, study 3 explored the moderating role of choice realism on goal satiation by examining the impact of a first choice that was either real or hypothetical on a second (real) choice.

Design and Procedure

The study used a 3 (goal prime: prestige vs. thrift vs. neutral) \times 2 (goal satiation: high [real choice] vs. low [hypothetical choice]) between-subjects design. In the neutral goal prime conditions, neutral words replaced the prestige or thrift words in the scrambled-sentence task. One hundred and eighty participants were randomly assigned to one of the six conditions.

The study was carried out in two different rooms. In the

first room, the goal prime factor was manipulated as in studies 1 and 2. Following the goal priming task, participants watched a video for 5 minutes. Participants then moved to the second room, one individual at a time. On their way to the second room, participants chose between one pair of Tommy Hilfiger socks against three pairs of Hanes (both options valued at \$6). Participants in the high-goal-satiation conditions read this statement: "This is a real choice. That is, you will actually receive the option you pick. Now, please check the option of your choice." Participants in the low-goal-satiation conditions saw: "Pretend that this is a real choice. That is, pretend that you will actually receive the option you pick." Participants in the neutral goal prime conditions also engaged in the intervening task. Their choices in both the intervening task and the final choice task were no different across the goal-satiation conditions. Their responses were, therefore, collapsed across the two goal-satiation conditions to create a control condition. As a consequence, our analyses were carried out with the following design: 2 (goal prime: prestige vs. thrift) \times 2 (goal satiation: high vs. low) plus a control.

Participants then entered another room, where they learned about the second choice through the following statement: "By taking part in this study, you will automatically be entered in a lucky draw. Two winners will be declared at the end of the experiment. If you are the winner, you will receive a prize worth \$100. Please let us know which of the two options on display you would like, should you be the winner of the lottery." Participants were then shown a display that featured (a) Timex watches (ladies' or men's), with a sign stating: "TIMEX (\$22.50 value) plus \$77.50 in cash," and (b) Guess watches, with an accompanying sign stating "GUESS (\$75 value) plus \$25 in cash." Participants indicated their choice on the instrument that was handed to them. Care was taken that the two choice tasks occurred completely out of sight and out of earshot of the other respon-

dents and the experimenters, in order to control for social and normative factors, which could have otherwise influenced participants' decisions (see Ratner and Kahn 2002; West and Broniarczyk 1998).

Finally, participants responded to a series of measures and were then debriefed. The measures included a manipulation check for the goal satiation factor and checks for potential confounds. Specifically, participants indicated on seven-point scales the extent to which they (*a*) treated the choice tasks as being real, (*b*) deliberated, (*c*) paid attention, (*d*) spent time in making up their minds, and (*e*) treated the task seriously.

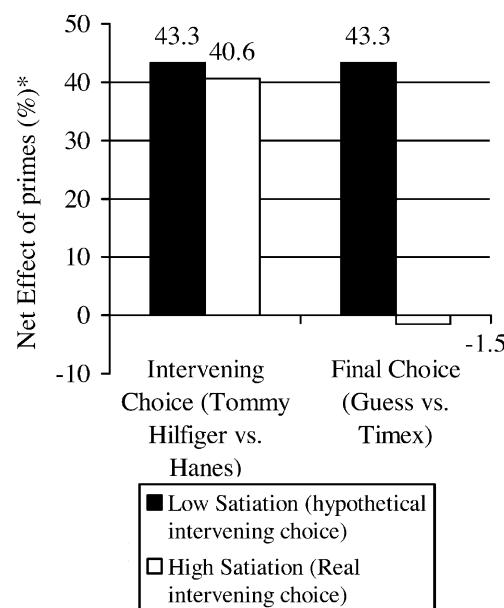
Results

Intervening Choice. A logit model on intervening choices revealed a significant main effect of the goal prime factor ($\chi^2 = 20.27, p < .0001$). None of the other treatment effects were significant. The pattern of results was in line with our conceptualization. Specifically, in the first choice task the selection of Tommy Hilfiger socks was higher in the prestige condition (62.8% and 62.2% in the high- and low-goal-satiation conditions, respectively) compared with the thrift goal prime condition (22.2% and 18.9% in the high- and low-goal-satiation conditions, respectively; $\chi^2 = 11.18, p < .0008$, in the high-goal-satiation conditions; $\chi^2 = 13.09, p < .0003$, in the low-goal-satiation conditions). The choice of Tommy Hilfiger socks in the control condition was 40%.

Effect of Satiation. More critical for our hypothesis, the goal-priming manipulation carried over to the subsequent "final" choice task when goal satiation was low (i.e., when the intervening choice was hypothetical) but not when it was high (i.e., when the intervening choice was real). Specifically, a logit model on the choices made in the final choice task revealed a significant goal prime by goal satiation interaction ($\chi^2 = 7.62, p < .006$), in addition to significant main effects of goal prime ($\chi^2 = 8.0, p < .02$) and goal satiation ($\chi^2 = 4.17, p < .04$). Within the high goal satiation conditions, choice of the Guess watch (plus \$25 in cash) was no different across levels of the goal prime factor (42.9% and 44.4%, in the prestige and thrift conditions, respectively; choice of the Guess watch in the control condition was 42.9%). In contrast, within the low goal satiation hypothetical conditions, choice of the Guess watch was higher in the prestige condition (64.9%) than in the thrift condition (21.6%; $\chi^2 = 13.0, p < .0003$).

Stated differently and as shown in figure 2, when participants had previously made a real choice, the net effect of priming (i.e., the difference in choice of the Guess watch between the prestige and the thrift prime conditions) vanished in the second choice (-1.5%) in contrast to prime-consistent behavior when the previous choice was hypothetical (43.3%; $\chi^2 = 7.62, p < .006$). The hypothetical results replicate what was found with hypothetical choices in study 2, while the real choices demonstrate extinction of the cued goals in line with the goal satiation criterion.

FIGURE 2
NET EFFECT OF PRIMES ON INTERVENING AND FINAL CHOICES: STUDY 3



Also of note, the results relative to the control display largely symmetric effects for the prestige and the thrift primes. That is to say, the percentage of participants choosing the more expensive options in the intervening and the final choice control conditions (40% and 43%, respectively) lies midway between the choice percentages for the appropriate prestige and thrift conditions (excluding the high goal satiation final choice condition, where the effect of the primes was attenuated).

Discussion

In addition to replicating the effect of subtle goal activation on consumer choice found in studies 1 and 2, study 3 provides further support for our motivational account by demonstrating attenuation of the effect of the prime after a real choice is made. We believe this provides the first empirical demonstration of the satiation of nonconsciously held and pursued goals. Furthermore, study 3 builds on study 2 by documenting the important finding that while real choices can satiate nonconsciously held goals, hypothetical choices may not.

The three studies provide convergent evidence that consumer choices are influenced by nonconscious goals operating outside of the consumers' awareness and intent. These studies rely on a supraliminal priming procedure (i.e., the scrambled-sentence task). We use subliminal priming—an even more conservative priming paradigm—in study 4 to provide convergent validity and increase the generalizability

of our findings. Moreover, by making the subliminal cues brand names, we test the possibility that brands may be a viable trigger for the activation of consumer goals. Study 4 addresses this issue by measuring preferences for prestige- and thrift-oriented products after subliminal presentation of names of retail stores.

STUDY 4: BRANDS AS SUBLIMINAL CUES TO GOAL ACTIVATION

Faced with the mounting evidence that goal priming can affect behavior outside of awareness and intent, research has started to explore naturalistic triggers of nonconscious goals. In the original conception of nonconscious goal pursuit (Bargh 1990), environmental features that are consistently paired with goals develop the ability to automatically activate those goals in the future. Thus, a social party environment may trigger nonconscious self-presentation goals in an individual chronically concerned with appearances. Researchers have identified a variety of naturalistic environmental triggers of nonconscious goal pursuit, including situations of power, self-esteem threat, presence of temptations, significant others, and situational norms (Aarts et al. 2005; Aarts and Dijksterhuis 2000; Bargh et al. 1995; Chartrand, Dalton, and Fitzsimons 2008; Fishbach et al. 2003; Fitzsimons and Bargh 2003; Shah 2003; Spencer et al. 1998).

Study 4 investigates an unexplored class of potential triggers of nonconscious goals—retail brands. As illustrated in our initial vignette, the potential for such brands to become cues for goal activation is of particular interest to marketers whose customers are exposed to brands throughout the entire purchase process. For example, when a consumer visits Wal-Mart with a goal to save money and sees its price-cutter ads, the goal of thrift may become linked to exposure to Wal-Mart. If so, a consumer on the way to a sports store who passes a Wal-Mart on the way may have a thrift goal nonconsciously activated that then may influence sock selection in the sports store. Automatic goal activation may thus open the door to an intriguing new way to think about the role and power of brands.

Design and Procedure

Study 4 randomly assigned 107 participants to a simple two-cell (subliminal prime: prestige retail brands vs. thrift retail brands) between-subjects design. The experiment was implemented in a room containing six personal computers with dividing screens between each to ensure that participants were not distracted by what others were doing. Participants were asked to remain upright in their chairs and to adjust the screen height to ensure its center was level with their eyes, so that the priming stimuli would be presented outside of participants' foveal visual field. Dimmed lights prior to the priming task enhanced the conditions for priming (Bargh and Chartrand 2000). The visual acuity exercise consisted of two tasks. First, respondents were asked to calculate and recall the cumulative total of numbers

flashed on the center of the screen. This task ensured that participants focused on the center of the screen. Second, participants were informed that a series of random flashes would occur throughout the experiment and were asked to indicate if the flashes occurred on the left or right side of the screen by pressing either the *f* key or the *k* key on their keyboard as soon as possible after they saw the flash.

Each flash consisted of the stimulus prime for 60 milliseconds immediately followed by a 100-millisecond backward mask (i.e., a string of letters in the same location). The masking string ("XQFBZRMQWGBX") contained a variety of letter patterns and was structurally similar to the preceding stimuli. To further ensure that participants focused on the center of the screen, they were told: "Because the location and timing of the flashes are random, it is important to keep your eyes on the center of the screen if you are to perform well on both tasks."

Each stimulus word appeared at one of four locations on the computer screen (i.e., in each of the four corners). At these locations the stimulus words were within participants' parafoveal visual field but outside the foveal visual field associated with conscious awareness (Bargh and Chartrand 2000). We constructed a randomized-location order and gave all participants the same sequence, with 2 seconds separating each stimulus word. The combination of these precautions—brief prime duration, immediate masking and placement in the parafoveal processing area—have been found to prevent awareness of priming stimuli in previous subliminal priming studies (e.g., Bargh and Chartrand 2000; Chartrand and Bargh 1996).

The stimulus words used were U.S. retail brands presented in Times New Roman typeface. In the prestige condition, the brands were Tiffany, Neiman Marcus, and Nordstrom, while in the thrift condition they were Wall-Mart, Kmart, and Dollar Store. These particular brands emerged from a pretest of 35 participants who listed retail brands that they considered to be strongly prestige oriented or thrift oriented.

Following the visual acuity task, participants indicated their relative preference of either a thrift or a prestige option for both a socks choice and a microwave choice. The combed-cotton socks scenario was identical to that used in studies 1 and 2; the Nike socks at \$5.25 for one pair, the Hanes socks at \$6.00 for two pairs. In the microwave scenario, participants were told: "Your microwave has broken, and you are considering the following two options." The available options were a 1-cubic-foot Haier microwave for \$69 or a 1-cubic-foot Sharp microwave for \$99. Preferences in each scenario were captured on a 1–7 scale anchored at the bottom, middle, and top with the terms "strong preference for [name of thrift option]," "indifferent," and "strong preference for [name of prestige option]."

Upon their arrival, the experimenter showed participants to their seats and told them that they would be taking part in two unrelated studies, the details of which were briefly explained. The supposed unrelatedness of the studies was reinforced by the fact that participants had seen faculty names from both the psychology and the marketing departments in

the electronic sign-up process. Participants took six practice trials to become familiar with the nature of the first task, each trial consisting of one number and one flash. The flashes in the practice trial consisted of random letter strings rather than brand names. The computer then informed them of the correct cumulative total for the practice trials and requested participants to press a key to start the actual trials. Participants then completed 75 experimental trials (25 exposures to each brand in a random order) in approximately 4 minutes.

After the participants had completed the priming procedure, the computer asked them to continue to the second experiment, in which they indicated their preferences for the two consumer product choices previously described. Finally, participants were thanked for their participation and debriefed.

Results and Discussion

The relative preference ratings of the prestige and the thrift options are shown in table 2. A repeated-measures ANOVA revealed a significant main effect of brand prime ($F(1, 105) = 3.58, p = .03$). As expected, for both the socks and the microwave scenarios, participants primed with the thrift-orientated brands expressed stronger relative preference for the thrift option than those participants primed with the prestige brands.

The responses to the funneled debriefing questions once again supported our contention that participants were unaware of the goal activation and the subsequent goal-driven nature of their behavior. None of the participants suggested that their choices in the second study had been influenced by the first. Even when subsequently prompted to guess at a possible link, participants did not come close to detecting the true design. We note that this continued absence of any reported awareness by participants of the primes' influence on their behavior, combined with the subliminal nature of the primes used in this study, renders possible demand explanations for our results extremely unlikely.

Thus, the subliminal exposure to retailers' names associated with either thrift or prestige products appeared to activate associated goals in our participants. To the best of our knowledge, this provides the first evidence that such brands can automatically activate purchase goals in individuals and that these goals can influence consumers' product preferences without their awareness or conscious intent.

GENERAL DISCUSSION

Taken together, the four studies demonstrate that fundamental consumer goals can be automatically activated and pursued outside of conscious awareness. Through the use of both supraliminal primes (studies 1, 2, and 3) and subliminal primes (study 4), as well as extensive debriefing, we establish that the goals are activated without the conscious knowledge of the participants, and we demonstrate that these goals then operate outside of awareness to influence choices and preferences. Empirical support for our motivational account includes both the observed strengthening of the choice

TABLE 2

RETAIL BRAND NAMES AS CUES TO GOAL ACTIVATION:
STUDY 4

	High-end store primes		Low-end store primes	
	Preference rating	SD	Preference rating	SD
Sock choice	3.22	2.1	2.60	1.9
Microwave choice	2.65	1.9	2.05	1.5

NOTE.—Higher numbers indicate greater preference for prestige option on 1–7 scale.

effects after a delay (study 2) and the attenuation of the effect of the prime after a real choice has been made and goal satiation has occurred (study 3). Study 4 additionally demonstrates that subliminal exposure to retail brands can trigger similar goal activation.

Implications for Consumer Decision Making

These results may seem surprising because the idea of nonconscious goals is so paradoxical. After all, goal pursuit seems almost by definition to be an intentional process. From a conceptual perspective, it is hard to imagine that a drive for a desired end state, with all the accompanying complexities associated with such goal striving (e.g., the need for flexible actions during goal pursuit in response to obstacles), can occur without conscious awareness. Further, to be effective, goal strength needs to vary; it needs to intensify initially but decrease when satiated. Nevertheless, our studies demonstrate that all this is indeed possible during nonconscious goal pursuit.

We may well prefer to believe that, as rational beings, we retain full conscious control of our actions, thoughtfully processing all the stimuli our senses collect and arriving at sensible conclusions and optimal decisions. However, our work contributes to a growing body of research that suggests that a substantial part of our adaptive mental functioning is rooted in nonconscious processing and thus challenges the rational model of decision making as a functionally impractical view. In our view, the ability to invoke goals automatically from environmental cues without substantial conscious processing is both efficient and adaptive and is no less relevant in consumer settings than in a number of other contexts. For example, consider the usefulness of automatically invoking goals to resist temptations when walking down the ice cream aisle or to avoid making a quick decision when dealing with a persuasive salesperson.

The current work suggests that the role of nonconscious processing may be as relevant to consumer choice as it is, for example, to memory, social perception, emotional appraisal, or causal attribution. Consumers may think they fully understand why they like and choose the products they do, but the current studies suggest there are at least three components of the process of which they can be unaware: (a) the activation by incidental environmental cues of goals outside of awareness, (b) the nonconscious pursuit of these

goals, and (c) the resulting consequences of these goals for choice or expressed preference (see also Chartrand 2005). Clearly, the importance of these issues is not restricted to marketplace choices alone. Researchers need to be aware of the potential consequences of nonconscious goal activation in their studies. Whenever the effects of an initial manipulation on subsequent behavior are examined, the possibility of the manipulation cuing unintended goal activation, in addition to affecting the construct of interest, needs to be considered.

Nor does nonconscious goal pursuit appear to be a fleeting or rare phenomenon. Our studies are consistent with prior research in demonstrating it is quite robust to the individual's potential awareness of the process—be it unambiguously nonconscious when subliminal primes are used (study 4) or nonconscious in the sense of awareness of the priming stimuli but not of the causal link (studies 1, 2, and 3). Furthermore, since study 4 demonstrates that retail names themselves can serve as the cues that activate such goals, and given the ubiquity of such brand exposure in daily life, our results are consistent with the emerging view that preferences may not be anywhere as stable as consumers would like to believe.

Implications for Nonconscious Goal Pursuit

The current results contribute to the nonconscious goal pursuit literature in two important ways. First, our work is the first to investigate the conditions under which nonconscious goals are satisfied and what happens when this goal satiation occurs. Study 3 demonstrates a clear attenuation of the effect of goal primes once a consequential goal-congruent choice has been made. While previous work has shown decreased goal activation upon fulfillment of consciously held goals (Forster et al. 2005), it has not been tested with goals operating outside of awareness. These findings are consistent with the emerging view that once activated, nonconsciously held goals operate in a fashion largely indistinguishable from that of their consciously pursued counterparts.

Second, our work provides the first evidence of the important finding that hypothetical goal-congruent choices are less satiating to nonconsciously held goals than are real choices. In study 3, we observed that changing the description of an otherwise identical choice from real to hypothetical was sufficient to attenuate goal satiation effects. An interesting question, one the current research did not address, is whether this property is unique to nonconsciously held goals. Would the conscious mind consider a hypothetical choice sufficient to fulfill a consciously held goal? We leave this question to future research.

This satiation finding joins a small but emerging body of work that identifies circumstances in which hypothetical choices may be imperfect surrogates for real choices. For example, the shopping-momentum effect (Dhar, Huber, and Khan, forthcoming) pertains to real but not hypothetical purchases. The current data suggest caution may sometimes be merited for studies utilizing multiple choices as depen-

dent variables. Researchers need to be aware of the fact that goal satiation effects have the potential to lead to disparities in choice between sequences of otherwise identical real and hypothetical options. While we continue to believe hypothetical choices are a valuable and efficient part of a consumer researcher's toolbox, the current research does raise a flag of caution, because—at least for satiation of nonconscious goals—they may not always be good proxies.

A second issue related to satiation concerns the specificity of the target goals. Our goals of thrift and prestige define directions rather than specific end states. Thus, it is not clear when either goal is satisfied and thus satiated. We expect that our results would be even stronger had there been a clearly defined end state to the elicited goals, such as a goal to spend less than a certain amount or to put together a high-status outfit. In those cases, one would expect the goal strength to increase as one approached the goal (e.g., Kivitz, Urminsky, and Zheng 2006) and would further expect a greater magnitude of goal satiation.

Marketplace Implications

It is important to recognize that our research utilized a tightly controlled laboratory environment in which participants were exposed to multiple congruent cues, all of which were designed to invoke the same goal. While our experimental environment may bear little resemblance to the chaotic display of images across stores at a mall, it is more consistent with the emphasis on coherent images within most stores. Consider the repetition of thrift images at Wal-Mart or the status images within Nordstrom. Still, it is important to understand how more heterogeneous goal cues in quick succession may activate unrelated, congruent, or opposing goals. Might initially activated goals suppress the activation of other goals until satiated? Is nonconscious goal activation an additive process, such that multiple cues of a similar type will lead to stronger behavior effects? Although now we may stand at only the beginning of the road to answering these questions, we believe there is clear potential for the current research to ultimately inform policy decisions.

First, given the potential complexities involved, we believe field experiments should focus on the clear managerial opportunities. For example, a retailer desirous of inducing consumers to trade up to higher-margin products in certain categories might want to try strategically embedding mini-displays of prestige items adjacent to displays of goods in those categories. Such "pockets of prestige" would aim to maximize the likelihood of prestige-type goals being activated in consumers as they browse the categories in question. Similarly, advertisements running on in-store displays (a growing trend) could be strategically biased toward prestige items in general.

Second, the time consumers spend in retail stores is often equal to or greater than the 8-minute period over which we observed substantial strengthening of goal-related behaviors. Thus, an additional factor for a retailer motivated to up sell consumers would be to expose shoppers to a disproportionate volume of prestige cues at the entrance to the

store and/or on end-of-aisle displays. The intent here would be to maximize the chance that prestige goals would be activated early in the shopping experience and to give them the maximum opportunity to strengthen before shoppers arrive at the parts of the store where their critical purchase decisions will be made. Appropriate positioning could increase the likelihood of high-margin choices while decreasing the probability that an intervening status purchase will satiate the goal.

An ideal context for testing the managerial implications of goal cuing arises out of actual or simulated retail Web sites. For instance, the banner side ads could be manipulated to evoke goals, with the placement of a focal promoted item either right after the cues or with a time lag. One could expect that the response to a promoted item will depend on the match between the cued goal and the focal product and that this response will initially increase with time but decrease with an intervening choice that satisfies that cued goal.

Directions and Implications for Future Research

Future research could profitably extend the current work in a number of different ways. First, although the current research was focused on examining the effects of priming thrift- and prestige-orientated goals, we believe that a wide variety of consumer goals can similarly be activated and pursued, influencing consumer choices in important ways. For example, to broaden the spectrum of goals examined, one might fruitfully investigate the nonconscious pursuit of effort and accuracy goals. Would, for instance, the nonconscious activation of an accuracy goal render a consumer less susceptible to falling prey to a positivity effect in inferring broad-service-provider quality from a single interaction (Folkes and Patrick 2003)? Widening the net of goals examined might also lead to research on regulatory orientation exhibited in prevention or promotion goals (Aaker and Lee 2006; Avnet and Higgins 2006) or on status versus functional choices (Poehlman et al. 2006).

Second, while the current research focuses on the effects of nonconscious goals on choice and preferences, future research might usefully explore subsequent downstream consequences such as product usage and enjoyment. For example, could consumers' enjoyment of a chosen option, or the attraction of a forgone option, be influenced by the goals that were activated at the time of purchase or use (Carmon, Wertenbroch, and Zeelenberg 2003; Nowlis, Mandel, and McCabe 2004)? Indeed, might perceptions of product performance, and thus of product satisfaction, be influenced by whether such a goal was consciously or nonconsciously activated and pursued (and, therefore, subsequently available or not available in memory)?

Third, more research is needed to examine the process by which nonconscious goals influence choice. For instance, is it possible that in a given choice task, nonconscious goals result in a biased weighting and/or evaluation of information as documented in research on predecisional distortion (Bond et al. 2007)? Finally, moving beyond research on goals per-

se, we believe the supraliminal and subliminal priming techniques used in the current research offer researchers a valuable tool with which to explore the moderation of given effects of interest. While other techniques often risk demand effects, the nonconscious nature of the priming techniques used here is largely immune to such concerns. For example, researchers interested in product contamination might use these priming techniques to temporally heighten accessibility of the disgust construct as part of an effort to examine its culpability as a moderator of perceived contamination (Argo, Dahl, and Morales 2006). The fact that nonconscious manipulations are, by definition, resistant to demand conditions, combined with the general finding that conscious and nonconscious goals act in very similar ways, suggests that research using nonconscious cuing can be expected to be more and more valuable in consumer research.

Conclusion

Despite the continued interest in the role of goals in consumer decision making, not much attention has been paid to factors that determine how goals actually come to be selected and pursued. Building on the notion that goals can be pursued nonconsciously (e.g., Bargh 2002; Chartrand et al., forthcoming), this article presents evidence suggesting that goals can be activated by situational cues and pursued until satiated by real choices and that the activation as well as the pursuit can occur nonconsciously. This article is, therefore, an attempt at challenging the notion that mental functioning is and needs to be conscious—a notion widely held among marketers—by presenting a case for revising it on the lines that a substantial part of our adaptive mental functioning may actually be rooted in nonconscious processing.

REFERENCES

- Aaker, Jennifer L. and Angela Y. Lee (2006), "Understanding Regulatory Fit," *Journal of Marketing Research*, 43 (February), 15–19.
- Aarts, Hank, Tanya L. Chartrand, Ruud Custers, Unna Danner, Giel Dik, Valerie E. Jefferis, and Clara Cheng (2005), "Social Stereotypes and Automatic Goal Pursuit," *Social Cognition*, 23 (December), 465–90.
- Aarts, Hank and Ap Dijksterhuis (2000), "Habits as Knowledge Structures: Automaticity in Goal-Directed Behavior," *Journal of Personality and Social Psychology*, 78 (January), 53–63.
- Argo, Jennifer J., Darren W. Dahl, and Andrea C. Morales (2006), "Consumer Contamination: How Consumers React to Products Touched by Others," *Journal of Marketing*, 70 (April), 81–94.
- Atkinson, John W. and David Birch (1970), *A Dynamic Theory of Action*, New York: Wiley.
- Avnet, Tamar and E. Tory Higgins (2006), "How Regulatory Fit Affects Value in Consumer Choices and Opinions," *Journal of Marketing Research*, 43 (February), 1–10.
- Bagozzi, Richard P. and Utpal Dholakia (1999), "Goal-Setting and Goal-Striving in Consumer Behavior," *Journal of Marketing*, 63 (Special Issue), 19–32.
- Bargh, John A. (1990), "Auto-motives: Preconscious Determinants

- of Social Interaction," in *Handbook of Motivation and Cognition*, Vol. 2, ed. E. Tory Higgins and Richard M. Sorrentino, New York: Guilford, 93–130.
- _____. (1997), "The Automaticity of Everyday Life," in *The Automaticity of Everyday Life: Advances in Social Cognition*, Vol. 10, ed. Robert S. Wyer Jr., Mahwah, NJ: Erlbaum, 1–61.
- _____. (2002), "Losing Consciousness: Automatic Influences on Consumer Judgment, Behavior, and Motivation," *Journal of Consumer Research*, 29 (September), 280–85.
- Bargh, John A. and Tanya L. Chartrand (2000), "Studying the Mind in the Middle: A Practical Guide to Priming and Automaticity Research," in *Handbook of Research Methods in Social Psychology*, ed. H. Reis and C. Judd, New York: Cambridge University Press, 253–85.
- Bargh, John A., Mark Chen, and Lara Burrows (1996), "Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action," *Journal of Personality and Social Psychology*, 71 (August), 230–44.
- Bargh, John A., Peter M. Gollwitzer, Annette Lee-Chai, Kimberly Barnsdollar, and Roman Trötschel (2001), "The Automated Will: Nonconscious Activation and Pursuit of Behavioral Goals," *Journal of Personality and Social Psychology*, 81 (December), 1014–27.
- Bargh, John A., Paula Raymond, John B. Pryor, and Fritz Strack (1995), "Attractiveness of the Underling: An Automatic Power-Sex Association and Its Consequences for Sexual Harassment and Aggression," *Journal of Personality and Social Psychology*, 68 (May), 768–81.
- Bettman, James R. (1979), "Memory Factors in Consumer Choice: A Review," *Journal of Marketing*, 43 (Spring), 37–53.
- Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (December), 187–217.
- Bond, Samuel D., Kurt A. Carlson, Margaret G. Meloy, J. Edward Russo, and Robin J. Tanner (2007), "Precommitment Bias in the Evaluation of a Single Option," *Organizational Behavior and Human Decision Processes*, 102 (March), 240–54.
- Carmon, Ziv, Klaus Wertenbroch, and Marcel Zeelenberg (2003), "Option Attachment: When Deliberating Makes Choosing Feel Like Losing," *Journal of Consumer Research*, 30 (June), 15–29.
- Chartrand, Tanya L. (2005), "The Role of Conscious Awareness in Consumer Behavior," *Journal of Consumer Psychology*, 15 (3), 203–10.
- Chartrand, Tanya L. and John A. Bargh (1996), "Automatic Activation of Impression Formation and Memorization Goals: Nonconscious Goal Priming Reproduces Effects of Explicit Task Instructions," *Journal of Personality and Social Psychology*, 71 (September), 464–78.
- _____. (2002), "Nonconscious Motivations: Their Operations, Activations, and Consequences," in *Self and Motivation: Emerging Psychological Perspectives*, ed. Abraham Tesser, Diederik Stapel, and Joanne Wood, Washington, DC: American Psychological Association, 13–41.
- Chartrand, Tanya L., Amy Dalton, and Clara Cheng (forthcoming), "Consequences of Nonconscious Goal Activation," in *Handbook of Motivation Science*, ed. James Shah and Wendy Gardner, New York: Guilford.
- Chartrand, Tanya L., Amy Dalton, and Gavan Fitzsimons (2008), "Nonconscious Relationship Reactance: When Significant Others Prime Opposing Goals," *Journal of Experimental Social Psychology*, 43 (September), 719–26.
- Custers, Ruud and Hank Aarts (2005), "Positive Affect as Implicit Motivator: On the Nonconscious Operation of Behavioral Goals," *Journal of Personality and Social Psychology*, 89 (August), 129–42.
- Dhar, Ravi, Joel Huber, and Uzma Khan (forthcoming), "The Shopping Momentum Effect," *Journal of Marketing Research*.
- Dijksterhuis, Ap and John A. Bargh (2001), "The Perception-Behavior Expressway: The Automatic Effects of Social Perception on Social Behavior," in *Advances in Experimental Social Psychology*, Vol. 33, ed. Mark P. Zanna, San Diego, CA: Academic Press, 1–40.
- Dijksterhuis, Ap, Tanya L. Chartrand, and Hank Aarts (forthcoming), "Automatic Behavior," in *Automatic Processes in Social Thinking and Behavior*, ed. John A. Bargh, Philadelphia: Psychology Press.
- Fishbach, Ayelet and Ravi Dhar (2005), "Goals as Excuses or Guides: The Liberating Effect of Perceived Goal Progress on Choice," *Journal of Consumer Research*, 32 (December), 370–77.
- Fishbach, Ayelet, Ronald S. Friedman, and Arie W. Kruglanski (2003), "Leading Us Not unto Temptation: Momentary Allurements Elicit Overriding Goal Activation," *Journal of Personality and Social Psychology*, 84 (February), 296–309.
- Fitzsimons, Gavan J. and Baba Shiv (2001), "Nonconscious and Contaminative Effects of Hypothetical Questions on Subsequent Decision Making," *Journal of Consumer Research*, 28 (September), 224–38.
- Fitzsimons, Gráinne M. and John A. Bargh (2003), "Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated with Relationship Partners," *Journal of Personality and Social Psychology*, 84 (January), 148–63.
- Folkes, Valerie S. and Vanessa M. Patrick (2003), "The Positivity Effect in Perceptions of Services: Seen One, Seen Them All?" *Journal of Consumer Research*, 30 (June), 125–36.
- Forster, Jens, Nira N. Lieberman, and E. Tory Higgins (2005), "Accessibility from Active and Fulfilled Goals," *Journal of Experimental Social Psychology*, 41 (May), 220–39.
- Hamilton, David L., Lawrence B. Katz, and Von O. Leirer (1980), "Cognitive Representation of Personality Impressions: Organizational Processes in First Impression Formation," *Journal of Personality and Social Psychology*, 39 (December), 1050–63.
- Higgins, E. Tory (2002), "How Self-Regulation Creates Distinct Values: The Case of Promotion and Prevention Decision Making," *Journal of Consumer Psychology*, 12 (3), 177–91.
- Huffman, Cynthia and Michael J. Houston (1993), "Goal-Oriented Experiences and the Development of Knowledge," *Journal of Consumer Research*, 20 (September), 190–207.
- Kivetz, Ran, Oleg Urminsky, and Yuhuang Zheng (2006), "The Goal-Gradient Hypothesis Resurrected, Purchase Acceleration, Illusionary Goal Progress and Customer Retention," *Journal of Marketing Research*, 43 (February), 39–58.
- Marsh, Richard L., Jason L. Hicks, and Martin L. Bink (1998), "Activation of Completed, Uncompleted, and Partially Completed Intentions," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24 (March), 350–61.
- Marsh, Richard L., Jason L. Hicks, and Eric S. Bryan (1999), "The Activation of Unrelated and Canceled Intentions," *Memory and Cognition*, 27 (March), 320–27.
- Nowlis, Stephen M., Naomi Mandel, and Deborah B. McCabe (2004), "The Effect of a Delay between Choice and Consumption on Consumption Enjoyment," *Journal of Consumer Research*, 31 (December), 502–9.
- Poehlman, T. Andrew, Eric Luis Uhlmann, Ravi Dhar, and John

- A. Bargh (2006), "Nonconscious Influences on Consumer Choice," Working paper, Marketing Department, Yale University.
- Ratner, Rebecca K. and Barbara E. Kahn (2002), "The Impact of Private versus Public Consumption on Variety-Seeking Behavior," *Journal of Consumer Research*, 29 (September), 246–57.
- Shah, James Y. (2003), "Automatic for the People: How Representations of Significant Others Implicitly Affect Goal Pursuit," *Journal of Personality and Social Psychology*, 84 (April), 661–81.
- Shah, James Y. and Arie W. Kruglanski (2003), "When Opportunity Knocks: Bottom-Up Priming of Goals by Means and Its Effects on Self-Regulation," *Journal of Personality and Social Psychology*, 84 (June), 1109–22.
- Shiv, Baba and Joel Huber (2000), "The Impact of Anticipating Satisfaction on Consumer Choice," *Journal of Consumer Research*, 27 (September), 202–16.
- Soman, Dilip and Amar Cheema (2004), "When Goals Are Counter-Productive: The Effects of Violating a Behavioral Goal on Performance," *Journal of Consumer Research*, 31 (June), 52–62.
- Spencer, Steven J., Steven Fein, Connie T. Wolfe, Christina Fong, and Meghan A. Dunn (1998), "Automatic Activation of Stereotypes: The Role of Self-Image Threat," *Personality and Social Psychology Bulletin*, 24 (November), 1139–52.
- Srull, Thomas K. and Robert S. Wyer (1979), "The Role of Category Accessibility in the Interpretation of Information about Persons: Some Determinants and Implications," *Journal of Personality and Social Psychology*, 37 (October), 1660–72.
- Van Osselaer, Stijn M. J., Suresh Ramanathan, Margaret Campbell, Joel B. Cohen, Jeannette K. Dale, Paul M. Herr, Chris Janiszewski, Arie Kruglanski, Angela Y. Lee, Stephen J. Read, J. Edward Russo, and Nader T. Tavassoli (2005), "Choice Based on Goals," *Marketing Letters*, 16 (December), 335–46.
- Watson, David, Lee Anna Clark, and Auke Tellegen (1988), "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales," *Journal of Personality and Social Psychology*, 54 (June), 1063–70.
- West, Patricia M. and Susan M. Broniarczyk (1998), "Integrating Multiple Opinions: The Role of Aspiration Level on Consumer Response to Critic Consensus," *Journal of Consumer Research*, 25 (June), 38–51.

In Defense of Consciousness: The Role of Conscious and Unconscious Inputs in Consumer Choice

Itamar Simonson

Stanford Graduate School of Business

Although the argument that unconscious inputs are often key determinants of consumer decision making is compelling, it may be overstated, particularly with respect to choice. A comparison of the effect of conscious inputs (e.g., the attributes of options in the choice set) and unconscious inputs (e.g., a seemingly irrelevant observation or task) indicates that the former have a significant advantage. In particular, the impact of conscious inputs is supported by choice task norms and is less susceptible to being lost in the “noise” that is characteristic of most natural consumer environments (e.g., stores). Indeed, although consumers often have limited insight into influences and processes producing their choices, the assumption that consumers base their choices on conscious, willful evaluation of task-relevant inputs has been quite successful in explaining a wide range of phenomena. It is expected that future research will put greater emphasis on the interactions between conscious and unconscious influences on decision making.

Building on a great deal of recent research, Dijksterhuis, Smith, van Baaren, and Wigboldus (2005) observed that “many choices are made unconsciously and are strongly affected by the environment” (p. 193). This conclusion is based on evidence concerning the link between perception and behavior, particularly the role of mimicry and activation of stereotypes, and on evidence regarding automatic goal pursuit. The notion that unconscious factors often have a significant effect on consumer choice and that such influences have been underresearched until recently is compelling. Furthermore, many of the studies that support the role of unconscious influences on judgment and behavior are noteworthy in their elegance and ingenuity, often demonstrating rather surprising effects. Thus, the article by Dijksterhuis and his colleagues (2005; see also Bargh, 2002) is likely to make an important contribution by raising consumer researchers’ awareness of the importance of focusing more attention on the ways in which unconscious, automatic processes might influence consumer decision making.

The conclusion that many psychological phenomena are largely determined by automatic, unconscious processes and inputs has received a great deal of attention in the past 15 years or so, though researchers have emphasized different aspects of such automatic, unconscious, intuitive effects (e.g., Bargh, 1997; Chaiken & Trope, 1999; Epstein, 1994;

Frederick, 2002; Kahneman, 2003; Sloman, 1996; Slovic, Finucane, Peters, & MacGregor, 2002). Integrating prior work, Kahneman (2003; see also Kahneman & Frederick, 2002) distinguished between the operations of *System 1* that tend to be automatic, effortless, associative, implicit, and often emotionally charged, and operations of *System 2* that are slower, consciously and deliberately monitored, and potentially rule governed. One of the conclusions that is implicitly or explicitly drawn from these related streams of research, including the work reviewed in the Dijksterhuis et al. (2005) article, is that conscious processes and consciously considered inputs play a relatively minor role in many, perhaps most, judgments, choices, and behaviors. For example, Bargh (1997) asserted that “everything that one encounters is preconsciously screened and classified as either good or bad, within a fraction of a second after encountering it” (p. 23). Loewenstein (2001) recently argued that “consciousness seems mainly to make sense of behavior after it is executed” (p. 503). Dijksterhuis (2004) argued that unconscious thought tends to improve the quality of decisions. Such notions go well beyond the well-established observation that decision makers often have limited insight into the determinants of their own judgments and decisions (e.g., Nisbett & Wilson, 1977). Furthermore, in making the distinction between conscious and unconscious influences, researchers have tended to paint the traditional emphasis on conscious inputs to decision making with a broad brush. Thus, critics of conscious decision research often point to studies that assumed that decisions are based

on (cognitive) evaluation of the various options' attributes or on a detailed listing of the options' pros and cons, as if conscious processing of perceived task-relevant inputs implies a comprehensive evaluation of all aspects.

More important, the proposition that automatic, unconscious influences are the primary drivers of decision making does not recognize significant distinctions within the very broad category of judgment, decision making, and behavior. As discussed in the following, the notion that automatic (System 1) influences are the default, with relatively infrequent override by conscious (System 2) processes (e.g., Bargh, 1997; Kahneman, 2003), may fit many psychological phenomena but does not adequately describe choice, where System 2 is usually the primary influence. In particular, (a) consciously considered inputs tend to play a major role in choice (including consumer choice), and (b) although understanding automatic, unconscious influences on choice is certainly important, the many potential unconscious influences in typical consumer-choice environments (e.g., in stores) create high "noise" level and potential interactions that tend to diminish the measurable significance of unconscious relative to conscious choice inputs.

ADVANTAGES OF CONSCIOUS INFLUENCES ON CONSUMER CHOICE

Dijksterhuis et al. (2005) motivated their discussion with an example of a shopper who finds himself at the supermarket counter with 26 items in the cart but cannot remember how most of them got there. Their explanation for the purchase of peanut butter, for example, is that "You hardly ever buy peanut butter, but a small boy running through the aisles reminded you of your 5-year-old nephew who loves peanut butter." Thus, seeing a small boy running through the aisles made the need for peanut butter salient, leading the consumer to put a jar of peanut butter in the cart. This example does not appear to depend primarily on unconscious influences and can be readily explained as a conscious process. It is now well accepted that consumers, and decision makers more generally, often construct their preferences when they need to decide, which makes them susceptible to a wide range of influences (for a review, see, e.g., Bettman, Luce, & Payne, 1998). Thus, contrary to the classical economic view of people's utility functions, it is no longer assumed that tastes are generally stable and well defined, and there is little doubt that the relative salience of decision criteria depends, among others, on the consumer's state of mind (see, e.g., Bettman & Sujan, 1987; Wright & Kriewall, 1980).

Returning to the peanut butter example, after seeing the boy, the mindless consumer evidently visited the supermarket section where peanut butter options were displayed. In all likelihood, given the lack of a compelling need for peanut butter, the decision to purchase peanut butter and the

act of placing it in the shopping cart involved at least some conscious processing of the observed stimuli. For example, if the nephew's favorite brand were not available, or if the store's price on that brand seemed unusually high, it would be quite possible that, the running boy's presence notwithstanding, peanut butter would not have been purchased. And even if the decision to purchase peanut butter was instantly made on observation of the child running in the aisles, it appears highly unlikely that the consumer would randomly select a peanut butter. Instead, the consumer would be likely to pay attention to one or more product attributes, such as his nephew's favorite brand, price, and fat content (given the sensitivity of the nephew's parents to that aspect). Overall, although the shopper would not recognize what triggered the peanut butter idea, the choice would involve a set of mostly conscious processes.

More generally, typical consumer-choice environments consist of the purchase options as well as many other stimuli. The former have a great advantage with respect to attention and impact on purchase decisions—they are the "main effects" that are usually perceived as relevant to the decisions to be made. Conversely, other inputs in the consumer environment (e.g., in stores) are usually not considered relevant, and their impact is more of an accident. Accidents and other low probability events do happen. For example, an observation of a hot dog or something that is yellow may very well make a consumer more receptive to purchasing an extra container of mustard. However, such effects, although clearly important to study and understand, are likely to have a smaller impact on typical consumers' choices than the consciously considered characteristics of the choice candidates and consumers' beliefs about their preferences.

The accessibility-diagnosticity framework of Feldman and Lynch (1988) addressed factors determining the likelihood that any cognition about an object will be used as an input to decisions concerning that or a related object. This framework can be used to assess the impact of conscious and unconscious inputs to consumer choice. According to this framework, the likelihood that any cognition will be used as an input is a function of (a) the accessibility of the input in memory, (b) the accessibility of alternative inputs, and (c) the diagnosticities of the input and of alternative inputs. The role and meaning of accessibility have been discussed extensively, although different definitions of accessibility have been applied (e.g., Higgins, 1996; Kahneman, 2003; Tulving & Pearlstone, 1966). For example, Kahneman (2003, p. 699) conceptualized accessibility broadly as determined by stimulus salience, selective attention, specific training, associative activation, and priming. According to Feldman and Lynch (1988), an input is diagnostic to the extent that consumers believe that the decision implied by that input alone would accomplish their decision goals (e.g., maximizing utility, choosing a justifiable option).

A comparison between conscious inputs to choice, particularly the characteristics of observed options, and uncon-

scious inputs in the consumer environment indicates that the former tend to have an overwhelming advantage on both the accessibility and diagnosticity dimensions. When making choices, it is customary to consider the options and their characteristics and make decisions accordingly. From a young age, children learn the ABCs of making choices and even become adaptive decision makers (e.g., Gregan-Paxton & Roedder John, 1995, 1997; Klayman, 1985). Although the level of choice involvement is often low and consumers' self-insight is rather limited (e.g., Wilson & Schooler, 1991), people generally believe that they should consider the options' characteristics when making decisions. This belief, in turn, implies that the characteristics of options tend to be more accessible and receive more attention than other, less directly relevant inputs. Moreover, the observed characteristics of options are generally perceived as diagnostic because people believe that they are the carriers of value and the proper bases for choice.

By contrast, the main "impact advantage" of unconscious inputs is that they are unconscious, making it unlikely that consumers would resist their influence. However, this factor seems much less significant when compared with the overwhelming disadvantage of unconscious inputs in terms of accessibility and perceived diagnosticity. Of course, consumers do not seek unconscious inputs that might influence their decisions, and they do not consider such inputs to be diagnostic or relevant. Indeed, had they been aware of the potential unconscious effects on their behavior, consumers would have likely tried and succeeded in eliminating them (although the degree of resistance to unconscious effects and the ability to control them might vary).

Another significant impact disadvantage of unconscious inputs is their high susceptibility to being lost in the "noise" that is characteristic of typical consumer (and many other) choice environments (e.g., stores, on the Internet). Because, unlike options' characteristics, potentially influential unconscious inputs are not sought by consumers, their status in the information processing hierarchy is not different from numerous other task-irrelevant inputs. For example, in addition to seeing the running boy, the shopper described by Dijksterhuis et al. (2005) might have also been exposed to an obese child, which could have negatively affected the likelihood of buying peanut butter. Dijksterhuis et al. reviewed work demonstrating the impact of (unconscious) mimicry on behavior (e.g., van Baaren, Holland, Steenaert, & van Knippenberg, 2003). But in a typical store environment, there are many people one might mimic. Can we predict or analyze with any accuracy such effects on consumer choice?

Although a great deal of progress has been made in recent years in our understanding of unconscious influences on judgment and decision making, there is little doubt that many other such effects are yet to be uncovered in this relatively new research area. As Bargh (1997) suggested, "It is hard to escape the forecast that as knowledge progresses re-

garding psychological phenomena, there will be less of a role played by the free will or conscious choice in accounting for them" (p. 1). We might discover, for example, that certain colors and shapes trigger particular responses and goals, and we may learn about various interactions among unconscious stimuli. In the analysis of consumer choice, such new insights would likely enhance the noise level, making it even more difficult to form predictions about the effects of unconscious inputs in natural consumer-choice environments. However, in an uncontrolled environment with many potentially significant sources of unconscious influence, predicting the overall effect of such inputs will be quite challenging.

The susceptibility of unconscious influences to being lost in the "noise" also has implications for the manner in which studies of consumer choice are conducted. In particular, whereas one can justify studying in isolation the impact of conscious inputs such as the characteristics of options and sets, isolating unconscious influences, although often intriguing and surprising, may not represent many real-world effects. Consider, for example, the research stream on context effects in choice (e.g., Huber, Payne, & Puto, 1982; Simonson & Tversky, 1992). Although these studies examined consumer response to very specific choice set configurations, the focal inputs represented the types of options that consumers actually focus on and evaluate in the process of choice (and, as shown by Kivetz, Netzer, & Srinivasan, 2004, these effects extend to more complex choice set configurations). That is, options represent the carriers of value that consumers usually intend to consider when making choices and, as a result, they often do consider such choice sets in the process of making decisions.

Conversely, unconscious influences are much less likely to operate in the clean form in which they are typically studied. That is, because such effects in real life are usually unintended coincidents, they are not protected by task goals and perceived relevance to the choice task. Any interference by other factors can eliminate or change the direction of such effects. For example, in the classic study by Bargh, Chen, and Burrows (1996), the presence of one participant who happened to be in a hurry at the conclusion of the study might have eliminated the effect of prior exposure to words related to the elderly on other participants' walking speed. Because most consumer-choice environments involve multiple potential unconscious influences, the likelihood that any one effect will operate in the isolated form in which it is demonstrated in the laboratory is relatively low. Furthermore, although measuring a choice (or other) dependent variable in close proximity to the manipulation is not unique to research on unconscious influences, it appears reasonable to assume that priming effects tend to deteriorate more rapidly than, for example, the effect of the considered options' attributes. As an aside, it is noteworthy that the conclusion that "everything that one encounters is preconsciously screened and classified as either good or bad, within a fraction of a second after en-

countering it"¹ (Bargh, 1997, p. 23; see also Duckworth, Bargh, Garcia, & Chaiken, 2001) may have limited consequences in many consumer-choice situations. In particular, preconscious classifications, affective reactions (e.g., Slovic et al., 2002), and confirmation bias notwithstanding, consumers often consciously consider attributes such as ingredients, features, and specifications, and these factors affect the ultimate choices they make.

CONSCIOUS CONSIDERATION OF TASK-RELEVANT INPUTS IN CONSUMER CHOICE

As Dijksterhuis et al. (2005) pointed out, the typical assumption underlying consumer research, and psychological research more generally, has been "that people consciously process information before they decide what to buy," whereas in reality they often do not. Consistent with this argument, Loewenstein (2001) suggested that "behavioral decision researchers are moving on" and "are abandoning their own paradigm" (pp. 499–500), and he reviews research indicating that decision making tends to be automatic, habitual, and mindless.

However, as indicated, at this point it appears highly unlikely that the explanatory power offered by an analysis of unconscious influences will approach that provided by the assumption that choices are largely determined by conscious processing of task-relevant inputs. This conclusion is particularly applicable to typical consumer choices and other choices, where conscious information processing is supported by both natural focal stimuli (i.e., the choice options) and by norms regarding the manner in which choices should be made. The conclusion may be somewhat less applicable to novel and unfamiliar judgment tasks, such as estimating distances or sizes, or determining how much one is willing to pay to save birds from drowning. But the assumption that consumers consciously consider the options available to them, whether these options fit their preferences, and so on, has been quite effective in allowing us to predict and explain a wide range of nonobvious marketplace phenomena. Of course, there is no one paradigm that can account for all choices and the manner in which information processing generates these choices. For example, whereas the simple assumption of value maximization can explain many observations, other phenomena appear to be better explained by portraying decisions as based on the balance of justifications for

and against options (e.g., Shafir, Simonson, & Tversky, 1993; Simonson, 1989; Slovic, 1975).

It must be emphasized again that the "conscious research program" has long abandoned the naive assumption that decision makers are aware of the various influences on their perceptions and behavior. Thus, a typical study of consumer decision making does assume that information processing of the manipulated stimuli or instructions or both takes place, but participants are often unaware of the factors driving their responses. The role of conscious information processing in the following illustrations, most of which are taken from projects in which I have been involved, is not unlike its apparent role in thousands of other studies. Although these examples are only briefly discussed, the assumption of conscious information processing of task-relevant inputs appears to account for these nonobvious influences on consumer-choice behavior.

Huber et al. (1982) demonstrated the attraction (or asymmetric dominance) effect, whereby the addition to a two-option set of an option that is inferior to one of the existing options increases the (absolute) choice share of that existing option. For example, consumers are more likely to exchange \$6 for an elegant Cross pen when they also have the option of exchanging \$6 for a less attractive pen (Simonson & Tversky, 1992). Evidently, the addition of the asymmetrically inferior option makes the superior option appear more attractive and easier to justify (e.g., Simonson, 1989), even though consumers fail to recognize the impact of the inferior option on their preferences (e.g., Dhar & Simonson, 2003). Thus, the robust asymmetric dominance effect appears to be driven by a rather detailed processing of the options' values and the set configuration, even though consumers tend to misattribute their choices to their tastes.

Kivetz and Simonson (2003) showed that idiosyncratic preferences, that is, preferences perceived to be different from those of most other people, play a key role and often receive disproportional weight in consumers' decisions. For example, they demonstrate that students who liked sushi more than most other students were more likely to join a loyalty program that offered a reward (movie tickets) for purchasing both 12 sandwiches and 12 orders of sushi than to join a program that offered the same reward for purchasing just 12 sandwiches. Kivetz and Simonson referred to the tendency to emphasize idiosyncratic preference fit as the *idiosyncratic fit heuristic*. Although people are likely to consciously process the provided information regarding aspects of options that fit their preferences, they do not recognize their tendency to emphasize such dimensions. For example, in a within-subjects version of this sushi study, respondents do not select the dominated program that required more purchases for the same reward.

Liu and Simonson (2004) asked one group of respondents to rank-order a set of rather similar See's chocolates. A second group was asked to rate the same chocolate options on a 0 to 100 scale. Next, participants in both groups were given a choice between \$2 and their favorite chocolate from the set.

¹It is important to note that such automatic classifications of stimuli as positive or negative are different from what we typically refer to as judgments or evaluations. The degree to which initial automatic classification of stimuli as good or bad determines the valence and intensity of conscious evaluations and the moderators of the relation between automatic classifications and conscious evaluations still need to be investigated (J. Bargh, personal communication).

The results indicated that those who rank ordered the options were significantly more likely to select their favorite chocolate over the \$2. This finding from a project in progress suggests that a ranking task produces preferences that are “closer” to choice. Again, although participants could not know what caused them to behave in a certain way, their choice was based on a conscious evaluation (i.e., ranking or rating) of the options. Finally, successful recent applications of framing by politicians and marketers have built on conscious processing of salient information as it is presented and on people’s failure to consider how they would have responded to alternative frames. Such examples include using the frames, (a) “climate change” rather than “global warming,” (b) “soy milk” rather than “soy juice,” (c) “death tax” rather than “estate tax,” and (d) “the Patriot Act,” “the Healthy Forest Initiative,” and “tax relief,” rather than alternative labels.

These examples are similar to numerous other illustrations of choice phenomena that can be readily explained based on the traditional assumption that choices are determined *primarily* by conscious, willful information processing of pertinent, task-relevant inputs, such as various interpretations of the options’ attributes and their fit with the person’s perceived preferences. As argued previously, because choices naturally focus on options, and people tend to believe that options need to be evaluated in some fashion before a choice is made, conscious accounts of choice behavior have a major advantage over unconscious influences.

The fact that all of the previously mentioned illustrations involve phenomena that can be characterized as primarily driven by conscious, willful, controllable evaluation of task-relevant focal inputs does not mean that decision makers are aware of the processes and the various factors (e.g., primes, goals, mood) that influence their responses. In that sense, one might argue these phenomena could also be regarded as unconscious and included under the “99 and $\frac{4}{100}\%$ ” of everyday life that is automatic (Bargh, 1997, p. 243). However, whereas the literature does not seem to offer a clear definition of automatic, unconscious influences (although Bargh, 1994, provided a conceptual classification), the emphasis and potential contribution of that literature go well beyond the well-established notion that people’s self-insight is limited. Indeed, observations of limited self-insight and failures of introspection have been well explained by analyses that have focused on these issues (e.g., Nisbett & Wilson, 1977; Wilson & Schooler, 1991).

CONCLUSION

The Dijksterhuis et al. (2005) article reviews and integrates recent research regarding unconscious, automatic influences on judgment, decision making, and behavior and will promote further research in this emerging, important area of consumer research. In this commentary, I tried to reflect

more broadly on arguments expressed in recent years by several prominent psychologists and decision researchers that unconscious, automatic influences are the primary drivers of judgment and choice. It is natural and often important for proponents of alternative views to highlight the common misconceptions and the underappreciation of their conviction.

However, some arguments are too extreme, such as the notion that conscious information processing of judgment and decision-making inputs (e.g., the observed options’ attribute values) usually just makes sense of behavior after it is executed. Similarly, the notion that unconscious, automatic processes determine most responses and account for 99 and $\frac{4}{100}\%$ appears overbroad and does not recognize important distinctions. In particular, the common assumption that choice is driven primarily by conscious processing of perceived task-relevant inputs still offers the most parsimonious account of choice behavior. Furthermore, although highly scripted or habitual responses might be considered nonconscious, they may be less interesting and reflect previously conscious processes.

Thus, it may not be meaningful to characterize judgments, decisions, and behavior as being normally nonconscious rather than conscious, or as System 1 rather than System 2. When discussing psychological phenomena that are driven mainly by automatic, unconscious processes, it seems reasonable to refer to consciousness and System 2 processes as overriding the default and automatic System 1 processes (e.g., Kahneman & Frederick, 2002; Sloman, 1996). Conversely, when accounting for choices and psychological phenomena that are driven mainly by task-relevant inputs, processes, and goals (e.g., attributes, tastes, rules), the characterization of System 2 as occasionally “overriding” System 1 seems less suitable.

The intriguing findings regarding unconscious, automatic influences on behavior do suggest promising directions for future research that incorporate both conscious and unconscious elements. In particular, using the types of manipulations described by Dijksterhuis et al. (2005), it should be possible to influence the criteria used by consumers (see, e.g., Bettman & Sujan, 1987) and the manner in which options are evaluated. For example, the tendency to consider regret and counterfactuals, to compromise, and to be in a frugal state of mind may very well be influenced by unconscious factors, such as imitation and goal pursuit.

Furthermore, reversing the typical order, consciously evaluated choice stimuli might affect performance in presumably unrelated priming procedures and, in turn, make consumers more susceptible to predictable priming effects. To illustrate, Dhar and Simonson (1999) showed that, in trade-offs between a goal (e.g., pleasure) and a resource (e.g., money) that involve two choices in the same episode, consumers tend to indicate a preference for “going all the way,” referred to as *highlighting*. For example, the same consumer is more likely to take the taxi to the airport when flying first class and more likely to take the

shuttle bus when flying coach. Conversely, in trade-offs between two goals, such as pleasure and good health, consumers tend to prefer balancing two components of an episode. For example, most consumers believe that they would be happier with two meals that balance a tasty but unhealthy appetizer/entrée with a less tasty but more healthy entrée/appetizer, as opposed to having an all-tasty meal on one occasion and an all-healthy meal on a second occasion.

Suppose, now, that study participants first consider two related choices and episodes (e.g., two dinners at a restaurant) involving a trade-off between pleasure and money (e.g., choices regarding the taste and cost of both the appetizer and entrée) and indicate a preference for total pleasure in one episode and low cost in the second episode (i.e., the highlighting rather than the balancing option). They are then asked to unscramble sentences, solve anagrams, or perform another seemingly unrelated task. Will participants who chose the highlighting option perform better in tasks where the solutions are compatible with the highlighted goal? Will such an effect be stronger for a goal (e.g., pleasure) than for a resource or a constraint (e.g., saving money)?

More generally, putting aside the question of whether conscious or unconscious, automatic aspects play a greater role in choice and can better explain consumer decision making, future research is likely to examine various interactions between conscious information processing and automatic influences and processes. Such research will promote the convergence of the two literatures. On the one hand, even the strongest supporters of consciousness recognize that unconscious, automatic processes can have significant impact on the manner in which consumers evaluate options and make choices. As the literature on automatic, unconscious influences further evolves, it is reasonable to expect that researchers will focus less on demonstrating that such effects exist and turn their attention to interactions between unconscious and conscious processes. Indeed, nonconscious influences may have their greatest and most enduring impact when they determine how decision makers consciously think about the objects of decision.

ACKNOWLEDGMENTS

This commentary has benefitted from the suggestions of Grainne Fitzsimons, Dale Griffin, Ran Kivetz, John Lynch, Stephen Nowlis, Lee Ross, Baba Shiv, and Christian Wheeler.

REFERENCES

- Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intentions, efficiency, and control. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., pp. 1–40), Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Bargh, J. A. (1997). The automaticity of everyday life, And, reply to commentaries. In R. S. Wyer Jr. (Ed.), *The automaticity of everyday life: Advances in social cognition* (Vol. 10, pp. 1–61). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Bargh, J. A. (2002). Losing consciousness: Automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research*, 29, 280–285.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). The automaticity of social behavior: Direct effects of trait concept and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244.
- Bettman, J., Luce, M. F., & Payne, J. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, 25, 187–217.
- Bettman, J., & Sujan, M. (1987). Effects of framing on evaluation of comparable and non-comparable alternatives by expert and novice consumers. *Journal of Consumer Research*, 14, 141–154.
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford.
- Dhar, R., & Simonson, I. (1999). Making complementary choices in consumption episodes: Highlighting versus balancing. *Journal of Marketing Research*, 36, 29–44.
- Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of Marketing Research*, 6, 146–160.
- Dijksterhuis, A. (2004). Think different: The merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology*, 87, 586–598.
- Dijksterhuis, A., Smith, P. K., van Baaren, R. B., & Wigboldus, D. H. J. (2005). The unconscious consumer: Effects of environment on consumer behavior. *Journal of Consumer Psychology*, 15, 193–202.
- Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2001). The automatic evaluation of novel stimuli. *Psychological Science*, 13, 513–519.
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist*, 49, 709–724.
- Feldman, J., & Lynch, J. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, 73, 421–435.
- Frederick, S. (2002). Automated choice heuristics. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 548–558). New York: Cambridge University Press.
- Gregan-Paxton, J., & Roedder John, D. (1995). Are young children adaptive decision makers? A study of age differences in information search behavior. *Journal of Consumer Research*, 21, 567–580.
- Gregan-Paxton, J., & Roedder John, D. (1997). The emergence of adaptive decision making in children. *Journal of Consumer Research*, 24, 43–56.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford.
- Huber, J., Payne, J. W., & Puto, C. P. (1982). Adding asymmetrically dominated alternatives: Violation of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9, 90–98.
- Kahneman, D. (2003). A perspective on judgment and choice. *American Psychologist*, 58, 697–720.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attributive substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 49–81). New York: Cambridge University Press.
- Kivetz, R., Netzer, O., & Srinivasan, V. (2004). Extending compromise effect models to complex buying situations and other context effects. *Journal of Marketing Research*, 41, 262–268.
- Kivetz, R., & Simonson, I. (2003). The idiosyncratic fit heuristic: Effort advantage as a determinant of consumer response to loyalty programs. *Journal of Marketing Research*, 40, 454–467.
- Klayman, J. (1985). Children's decision strategies and their adaptation to task characteristics. *Organizational Behavior and Human Decision Processes*, 35, 179–201.
- Liu, W., & Simonson, I. (2004). *Take it or leave it? Evaluation procedure as a moderator of the effect of preference on purchase*. Presentation at the Association of Consumer Research Conference, Portland, Oregon.

- Loewenstein, G. (2001). The creative destruction of decision research. *Journal of Consumer Research*, 28, 499–505.
- Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, 49(12), 11–36.
- Simonson, I. (1989). Choice based on reason: The case of attraction and compromise effects. *Journal of Consumer Research*, 16, 158–174.
- Simonson, I., & Tversky, A. (1992). Choice in context: Trade-off contrast and extremeness aversion. *Journal of Marketing Research*, 29, 281–295.
- Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119(1), 3–22.
- Slovic, P. (1975). Choice between equally valued alternatives. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 280–287.
- Slovic, P., Finucane, M., Peters, E., & MacGregor, D. (2002). The affect heuristic. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 397–420). New York: Cambridge University Press.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391.
- van Baaren, R. B., Holland, R.W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, 39, 393–398.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much—Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, 60, 181–192.
- Wright, P., & Kriewall, M. A. (1980). State-of-mind effects on the accuracy with which utility functions predict marketplace choice. *Journal of Marketing Research*, 17, 277–293.

Received: December 8, 2004



Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Consumer Psychology 22 (2012) 67–74

**Journal of
CONSUMER
PSYCHOLOGY**

Relative visual saliency differences induce sizable bias in consumer choice

Milica Milosavljevic ^{a,*}, Vidhya Navalpakkam ^a, Christof Koch ^a, Antonio Rangel ^{a,b}

^a Division of Biology, California Institute of Technology, MC 228-77, 1200 E. California Blvd, Pasadena, CA 91125, USA

^b Division of Humanities and Social Sciences, California Institute of Technology, MC 228-77, 1200 E. California Blvd, Pasadena, CA 91125, USA

Received 18 January 2011; received in revised form 1 September 2011; accepted 27 October 2011

Available online 3 December 2011

Abstract

Consumers often need to make very rapid choices among multiple brands (e.g., at a supermarket shelf) that differ both in their reward value (e.g., taste) and in their visual properties (e.g., color and brightness of the packaging). Since the visual properties of stimuli are known to influence visual attention, and attention is known to influence choices, this gives rise to a potential visual saliency bias in choices. We utilize experimental design from visual neuroscience in three real food choice experiments to measure the size of the visual saliency bias and how it changes with decision speed and cognitive load. Our results show that at rapid decision speeds visual saliency influences choices more than preferences do, that the bias increases with cognitive load, and that it is particularly strong when individuals do not have strong preferences among the options.

© 2011 Society for Consumer Psychology. Published by Elsevier Inc. All rights reserved.

Keywords: Decision making; Reward; Visual saliency; Attention; Packaging; Consumer choice

Introduction

Over the last three decades psychologists and behavioral economists have documented a large number of ways in which individual judgments and choices depart from optimal decision making and information processing (Gilovich, Griffin, & Kahneman, 2002; Kahneman & Tversky, 2000; Bettman, 1979). Most of these departures from rationality, often called biases, are due to a limited capacity for processing information that is relevant to the decision problem (e.g., memories about previous experiences, attributes of the choice stimuli, or the probabilities of potential gains and losses). This often leads to overweighting some relevant variables at the expense of others.

Recent work in visual and decision neuroscience suggests that the way in which the brain processes low-level visual and value information might also lead to decision making biases (Krajbich & Rangel, 2011; Krajbich, Armel, & Rangel, 2010; Rangel, Camerer, & Montague, 2008; Shimojo, Simion, Shimojo, & Scheier, 2003). The logic for this new class of

decision biases comes in two steps. First, a large body of work in visual neuroscience has shown that visual attributes of stimuli that affect their visual saliency, such as brightness or color, can affect the location and duration of fixations when individuals approach complex displays such as a vending machine or a supermarket shelf (Itti & Koch, 2001; Mannan, Kennard, & Husain, 2009). This visual saliency effect has been shown to persist for several fixations (Henderson, Weeks, & Hollingworth, 1999; Parkhurst, Law, & Niebur, 2002). As a result, more salient items are fixated on (i.e., looked at) longer than less salient stimuli. Second, a recent series of neuroeconomic studies have shown that the values assigned to stimuli at the time of choice depend on the amount of attention that they receive during the decision making process (Krajbich et al., 2010; Armel, Beaumel, & Rangel, 2008; Armel & Rangel, 2008; Shimojo et al., 2003). In particular, appetitive items receive higher liking ratings and are more likely to be chosen when attention focuses on them longer. Together, these two classes of findings suggest that everyday choices should be subject to visual saliency biases: independent of

* Corresponding author. Fax: +1 626 796 8876.

E-mail address: mmilosav@hss.caltech.edu (M. Milosavljevic).

consumers' preferences, more visually salient options are more likely to be chosen due to the specific way in which the brain processes visual information.

We present the results of three real food choice experiments designed to resemble every-day supermarket choices. In Experiment 1 we test for the presence of a *visual saliency bias* and measure its magnitude as a function of the length of exposure to choice alternatives, i.e., at varying speeds of decision making. In Experiment 2 we examine how the bias is affected by conditions of cognitive load pervasive in every-day life, such as talking on a cell-phone while shopping. Finally, in Experiment 3 we seek to provide external validity for the visual saliency bias by utilizing button-press choices, which are similar to every-day vending machine button-press purchases.

Experiment 1: visual saliency bias in every-day consumer choices

While academics and practitioners alike acknowledge that physical design of packaging influences purchases, no systematic knowledge exists of why some visual features "work" and others don't. Further, we know little about the relationship between visual features and preferences, and how this relationship changes over time. This first experiment was designed to quantify the impact that exogenously controlled differences in visual saliency, in this case brightness of the food items, have on every-day food choices and how this effect changes over time.

Methods

We here introduce a novel experimental paradigm from visual psychophysics and neuroscience. Seven "...participants completed..." the experiment. They were asked not to eat for 3 h prior to the experiment, which had two parts. In the first part participants indicated their food preferences by ranking 15 snack food items (such as chips and candy bars) from 1 (most favorite) to 15 (least favorite). The rankings were used as a measure of subjective value for each item and participant.

In the second part, each participant made 1050 choices between pairs of foods (Fig. 1). Each trial began with an enforced 800 ms central fixation. Next, a blank screen was flashed for 200 ms. This is thought to "reset" the visual system, by clearing any latent information, and to increase the speed of visual processing (Fischer & Weber, 1993). Participants were then shown two different items, one on the left and the other on the right side of the screen. In order to simulate a crowded display, such as a vending machine, each item was surrounded by 8 other items. The experiment included 5 trials for each possible item pair and left-right spatial configuration.

We experimentally manipulated relative preferences and relative visual saliency as follows. The absolute difference in liking rankings between the two items (d) provides a measure of the strength of preference, where $d=1$ represents weak preference and thus difficult choice, while $d=14$ represents strong preference and thus easiest choice. The visual saliency of the two food items was manipulated by changing their relative local brightness (see Fig. 1). On each trial one of the two

food items was made salient by decreasing the brightness of all other items to 65% (both the items surrounding the salient item and all items on the other side of the screen), while keeping the salient item's brightness at 100% so that it would visually "pop out"; i.e., be brighter than all other items.

To study how the visual saliency bias evolves with exposure time, food stimuli were displayed on the screen for 70, 100, 200, 300, or 500 ms. Presentation durations were kept constant in blocks of 105 trials. The block order was randomized across participants, and participants were given a short break between blocks.

Following the presentation of food items, a visual mask covered all the items (see Fig. 1). Visual masking is used frequently in vision science to stop any further visual processing, which allowed us to control the length of the exposure to the food items. The mask also signaled to participants that they should indicate their choice as quickly as possible by making an eye-movement towards the location of their preferred food item. At the end of the study, participants were asked to stay in the lab and were given to eat the food item that they chose in a randomly selected trial.

Gaze location was acquired from the right eye at 1000 Hz using an infrared Eyelink 1000 eyetracker (SR Research, Osgoode, Canada). Participants' heads were positioned in a forehead and chin rest. The distance between computer screen and participant was 80 cm, giving a total visual angle of $28^\circ \times 21^\circ$. The images were presented on a computer monitor using Matlab Psychophysics toolbox and Eyelink toolbox extensions (Brainard, 1997; Cornelissen, Peters, & Palmer, 2002). Left or right food choices were determined when an eye-movement was initiated and crossed a threshold of 2.2° from the center of the screen toward the left or the right food item. Response times are measured as the time from stimuli offset to response initiation.

About the experimental design

Several features of the experimental design, which are not common in the consumer research literature, are worth emphasizing. This is not a traditional eye-tracking study. In particular, we do not collect and analyze sequences of gaze data since the only recorded eye movements are those used to indicate choices, at which time the trial ends. Despite this, the eyetracker plays a critical role in the experimental design.

First, it is used to enforce an 800 ms fixation to the central fixation cross at the beginning of every trial. In particular, the choice alternatives are not displayed on the screen until fixation at the center of the screen has been verified by the eye-tracker. This is critical because visual saliency depends on the relative position of the subsequent stimuli with respect to the current location of fixation. Thus, if participants had been allowed to move their eyes freely during this time, we would have lost experimental control over our relative saliency manipulation.

Second, using the eye-tracker to record choices allows us to collect less variable measures of decision time. The intuition for why this is the case is simple. Standard measures of decision times, such as response times measured by a button press, are

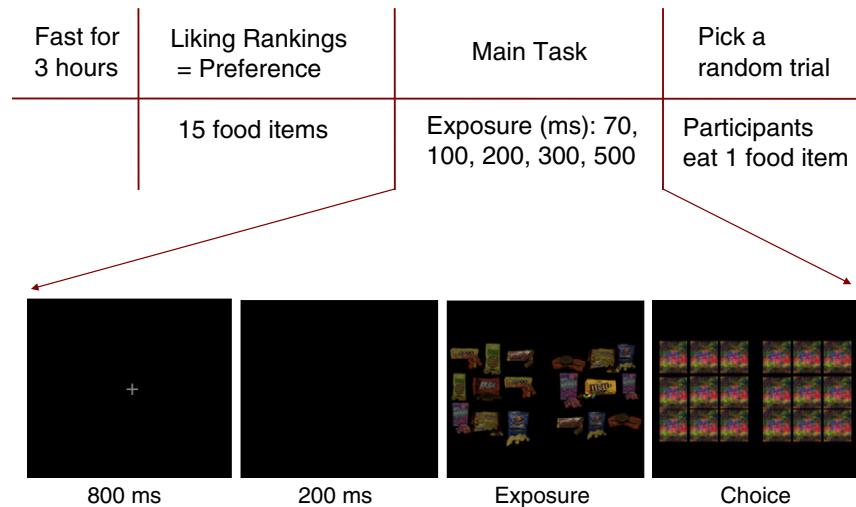


Fig. 1. Experiment 1. Hungry participants were asked to quickly decide whether they want to eat the left or the right central food item and to indicate their choices by looking toward the side of the screen where the chosen item was displayed. The saliency of one of the central items was enhanced by making it brighter, which caused it to “pop out” on the screen.

made up of two components: the time to decide and the time that it takes to implement the motor action used to indicate the choice. As a result, they have two sources of variability. By recording a decision as soon as an eye movement has been initiated, we can improve the precision of our measures of decision times since any variability of the time that it takes to execute the motor plan is eliminated.

Finally, we collect data on a very large number of trials within participants, but only for a handful of participants, which is standard practice in visual psychophysics and neuroscience. The rationale is simple: this allows the researchers to characterize the value of the parameters for each individual participant with great precision. Furthermore, a small number of participants are sufficient because there is often limited variation on the estimated properties of the underlying systems across participants.

Results

Mean response times by display duration were 317 ± 31 ms (70 ms exposure), 348 ± 43 ms (100 ms exposure), 281 ± 35 ms (200 ms exposure), 306 ± 37 ms (300 ms exposure), and 226 ± 10 ms (500 ms exposure). A repeated measures ANOVA did not reveal a significant effect of display duration on response time ($F=1.56$, $p>.22$).

Fig. 2a shows the percent of correct choices (i.e., how often participants chose their preferred food items as indicated by reported rankings at the beginning of the experiment) as a function of (1) exposure time shown on the X-axis, (2) strength of preference (i.e., the difference in ranking: strong=high brand dominance if $d \geq 7$ and weak=low brand dominance if $d < 7$), and (3) the presence of conflict between preference and visual saliency (i.e., conflict occurs when one item is more liked, but the other item is visually more salient). In the easiest condition (i.e., strong preference, high saliency of the preferred item, and longest exposure duration) participants chose their preferred

items 90.4% of time. A 3-way repeated measures ANOVA with no interaction effects showed significant effects of:

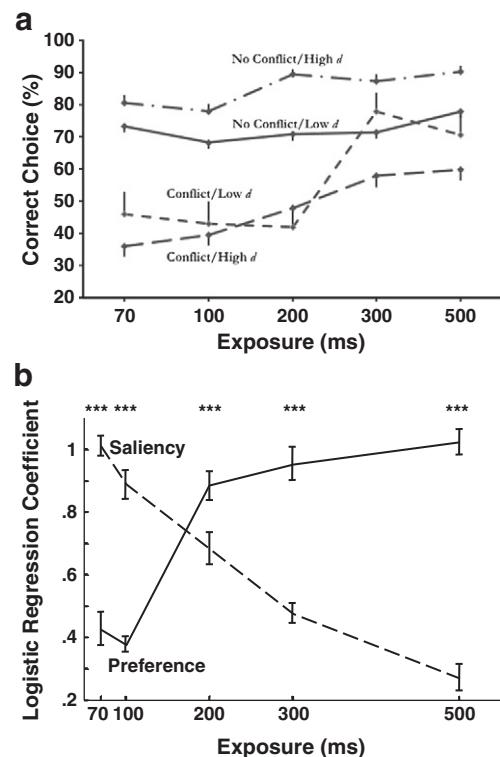


Fig. 2. Visual saliency bias in everyday consumer choice. (a) The percent of correct choices as a function of exposure, strength of preference, and whether or not there was a conflict between the preference and visual saliency. “No Conflict” denotes trials where the preferred food item is also more salient. “Conflict (high d)” includes trials where one of the two items is more liked but the other is more salient, and there is a strong prior preference for one of the items. “Conflict (low d)” refers to conflict trials where the strength of preference for one of the items is small. Error bars show S.E.M. (b) Estimated logistic regression coefficients for the relative preference and visual saliency as a function of exposure time. Error bars represent S.E.M. (**p<.001).

exposure time ($F=10.09$, $p<.00001$), strength of preference ($F=72.85$, $p<.00001$), and presence or absence of conflict between saliency and preference ($F=265.64$, $p<.00001$). Thus, participants were better at choosing their preferred food items as exposure time increased, when they had strong rather than weak preference for one of the two items, and when their preferred food items were also visually more prominent.

Next, we assessed the effects of preference and saliency on choice and how this relationship changes over time. We performed a mixed effects logistic regression of the probability of choosing the left item on its relative liking ranking (1 if left item is preferred and 0 otherwise) and its relative visual saliency (1 if left item is more salient and 0 otherwise). Fig. 2b shows the estimated coefficients for the relative preference and saliency variables as a function of exposure length (the regression was estimated separately at each exposure length). The influence of saliency on choice is dominant early on, at exposures of <200 ms ($p<.00001$, two-tailed t-test), but preference becomes the dominant influence after 200 ms ($p<.00001$, two-tailed t-test). Critically, however, saliency had a significant positive effect on choices at all exposure durations, even at 500 ms ($p<.02$; two-tailed t-test).

Finally, we examined how the influence of saliency vs. preference on choice depends on the underlying strength of preference. Fig. 3 depicts the choice patterns for conflict trials, in which one item is more liked but the other is more salient. When consumers have strong prior preference for a particular brand (Fig. 3a), they are quite good at identifying and choosing that brand rapidly: starting at 300 ms exposures, consumers are already choosing their preferred items over 70% of time even when these items are not visually prominent and are actually quite difficult to spot. However, in situations of brand parity (Fig. 3b), when the available brands are all pretty similar, consumers are influenced much more by visual features of the food items and end up choosing items that are visually prominent at least 40% of time even though this is inconsistent with their prior preferences.

Discussion

The results of Experiment 1 demonstrate a significant *visual saliency bias* in choices, which is dominant at very fast exposure times and persists even at longer exposure times. For example, the results show that at the shortest exposure times, visual saliency had over 200% more impact on choices than a 1-point increase in relative rankings, and the effect was still around 25% for the longest durations. Further, the bias is especially strong when participants' relative preferences between the two items are relatively weak, as is likely to be the case in many real world choices where market competition generates large numbers of similarly attractive options.

Experiment 2: visual saliency bias under cognitive load

Many every-day choices are made while decision-makers are engaged in alternative cognitively demanding tasks. For example, while purchasing a bag of chips in a grocery store, it is

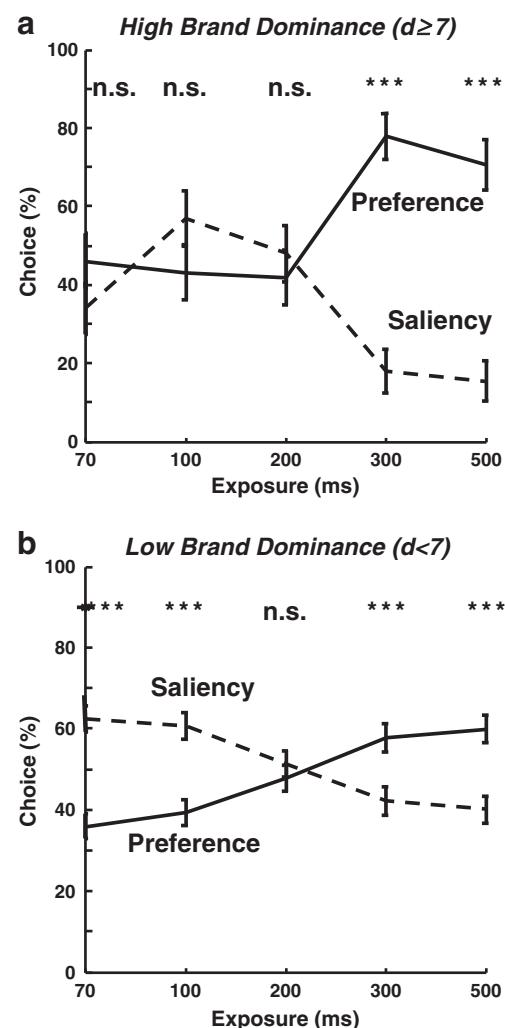


Fig. 3. Percent of choices on conflict trials (where one item is preferred but the other is more salient) for (a) easy choices with strong preferences for one of the items ($d \geq 7$) and (b) difficult choices with weak preference ($d < 7$). Error bars represent S.E.M. (**p<.001).

very likely that the consumer will also engage in additional tasks, such as a phone conversation or attending to a child that came along for the trip. Drolet, Luce, and Simonson (2009) recently suggested that such conditions of cognitive load do not necessarily make our choices worse, but rather increase the importance of externally available information. This provides motivation for our next experiment, since it suggests that visual saliency bias might become even more important under conditions of cognitive load. Experiment 2 examined this issue by introducing cognitive load into the previous experimental design.

Methods

Five "...participants completed..." this experiment, which is very similar to the previous one, except for the addition of a cognitive load task. On every trial the central fixation cross changed its color to either red or green. Participants were instructed to start each block of trials with a score of 50 and

add 1 to the score each time the cross turned green, and subtract 1 each time it turned red. As before, they were also asked to make a choice as soon as the visual mask appeared on the screen by moving their eyes to the preferred item's location. To accommodate the added cognitive demands of the task, items were now displayed at 100, 300, 500, 1000 and 1500 ms.

Results

Mean response times were 468 ± 50 ms (100 ms exposure), 371 ± 40 ms (300 ms exposure), 282 ± 25 ms (500 ms exposure), 292 ± 27 ms (1 second exposure), and 279 ± 9 ms (1.5 seconds exposure). A repeated measures ANOVA revealed a significant effect of display duration on response time ($F=4.86$, $p<.01$). Post-hoc multiple comparisons performed using Tukey's HSD test show that choices after 100 ms exposure were significantly slower (468 ± 50 ms) than those at 500 ms (282 ± 25 ms), 1000 ms (292 ± 27 ms) and 1500 ms (279 ± 9 ms) exposures.

All participants were able to perform the cognitive overload task, with performance improving as the length of exposure increased. The differences between participants' responses and correct answers (participants kept a count starting at 50 and could end as low as 30 or as high as 70) were 5.2 for 100 ms,

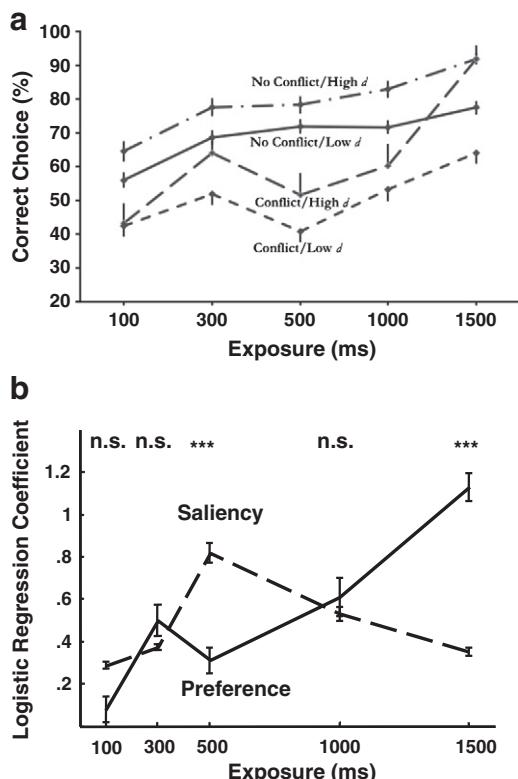


Fig. 4. Visual saliency bias under cognitive load. (a) The percent of correct choices as a function of exposure, strength of preference, and whether or not there was a conflict between the preference and visual saliency. Error bars show S.E.M. (b) Estimated logistic regression coefficients for relative preference and visual saliency as a function of exposure time. Error bars represent S.E.M. (***) $p<.001$.

2.1 for 300 ms, 2.2 for 500 ms, 1.7 for 1000 ms, and 0.8 for 1500 ms (repeated-measures ANOVA, $p=.056$).

Fig. 4a shows the percent of correct choices as a function of the length of exposure, strength of preference, and saliency vs. preference conflict. In the easiest condition (i.e., strong preference, high saliency of the preferred item, and longest exposure duration) participants chose their preferred items 92.0% of time. A 3-way repeated measures ANOVA with no interactions showed significant effects of: exposure duration ($F=36.98$, $p<.00001$), strength of preference, i.e., choice difficulty ($F=51.45$, $p<.00001$), and presence or absence of saliency-value conflict ($F=165.83$, $p<.00001$), all in the expected directions.

Fig. 4b depicts the estimated logistic regression coefficients for relative saliency and preference as a function of the exposure duration. The influence of preference on choice peaked later than without cognitive load (at 1500 ms here vs. 200 ms in Experiment 1; $p<.00001$, two-tailed t-test). Critically, the influence of saliency on choice remained strong at all exposure durations ($p<.00002$; two-tailed t-test).

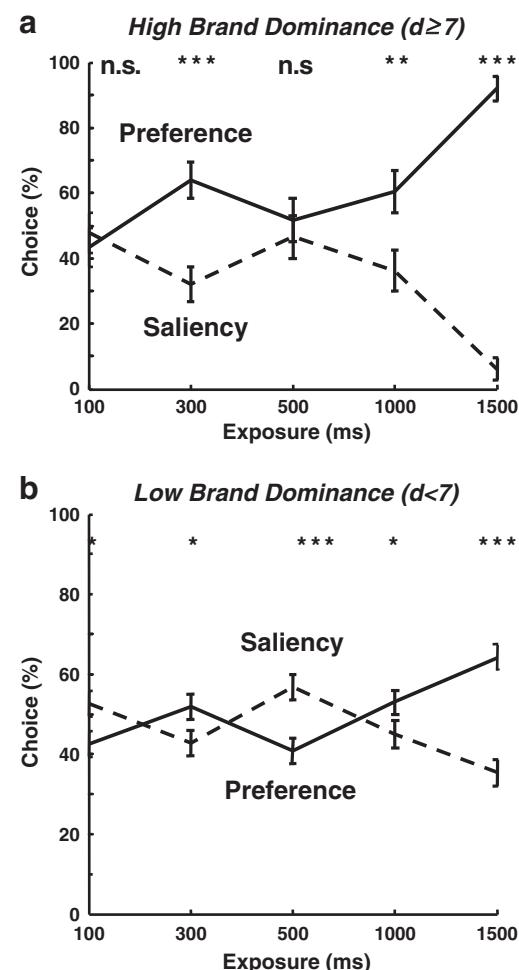


Fig. 5. Percent of choices on conflict trials (where one item is preferred but the other is more salient) for (a) easy choices with strong preferences for one of the items ($d \geq 7$) and (b) difficult choices with weak preference ($d < 7$). Error bars represent S.E.M. (***) $p<.001$.

Finally, we examined how the influence of saliency vs. preference on choice depends on the strength of preference. Fig. 5 shows conflict trials, where one item is more liked but the other is more salient. Only once the exposure duration reached 1.5 s did preference clearly take over, and it did much more so for high brand dominance (Fig. 5a) than in the condition of low brand dominance (Fig. 5b).

Discussion

The results of Experiment 2 show that the visual saliency bias is longer-lasting, and relatively stronger, in the presence of cognitive load. For example, at stimuli exposures of 500 ms, the impact of visual saliency was over 200% stronger than that of a 1-point increase in the liking rankings. Importantly, the visual saliency bias was still significant and sizable at exposure times as long as 1500 ms.

One striking aspect of the data, which we were surprised by, was that the effect of saliency on choice did not decrease monotonically with exposure time (see Fig. 4b). We have the following post-hoc explanation for why this is the case. In previous work (Krajbich et al., 2010) we have shown that these types of choices seem to be made through a temporal integration model that looks like an attention modulated version of the Drift-Diffusion-Model (Ratcliff & Smith, 2004; Ratcliff & McKoon, 2008). One critical feature of these models is that early on in the integration process there is a lot of noise in the identification of the best item. As a result, early on, the choices are not significantly related to any of the underlying variables driving the computation and comparison of values. From this perspective, the data for Experiments 2 and 3 (discussed next) suggests that the cognitive load manipulation is slowing down the comparator process, and thus the effect of both saliency and preference is showing up later than in Experiment 1.

Experiment 3: external validity of visual saliency bias under cognitive load

One concern with the previous two experiments is that participants indicated their choices with eye movements. Experiment 3 extends the results to the case of hand movements, another motor modality that is representative of every-day behavior (e.g. a button press during vending-machine purchases).

Methods

The study follows the same procedure as Experiment 2, except that now participants were asked to indicate their choices by pressing the left or the right arrow key on the keyboard to indicate the location of the preferred item on the screen. The same five participants from Experiment 2 completed this task.

Results

Mean response times were 498 ± 11 ms (100 ms exposure), 343 ± 8 ms (300 ms exposure), 299 ± 7 ms (500 ms exposure), 296 ± 10 ms (1000 ms exposure), and 321 ± 9 ms (1500 ms

exposure). A repeated measures ANOVA revealed a significant effect of exposure duration on time that it took to indicate choices ($F=4.49$, $p<.01$), but a post-hoc Tukey's HSD test shows that this result is driven by the fact that the fastest exposure duration (100 ms) is significantly different from the three other conditions (500 ms, $p<.05$; 1000 ms, $p<.05$; and 1500 ms, $p=.05$). None of the other comparisons was significantly different from one another.

The difference between participants' responses and correct answers for the cognitive overload task shows that participants were able to perform the task well in all conditions (per block score: 1.7 for 100 ms, 2.5 for 300 ms, 3.1 for 500 ms, 2 for 1000 ms, and 1.3 for 1500 ms; repeated-measures ANOVA, $p=.581$).

Fig. 6a shows the percent of correct choices as a function of exposure time, strength of preference, and presence of the saliency vs. preference conflict. In the easiest condition (i.e., strong preference, high saliency of the preferred item, and longest exposure duration) participants chose their preferred items 92.7% of time. As in previous experiments, a 3-way repeated measures ANOVA with no interactions showed that the three variables had a significant impact on accuracy: exposure duration ($F=45.23$, $p<.00001$), strength

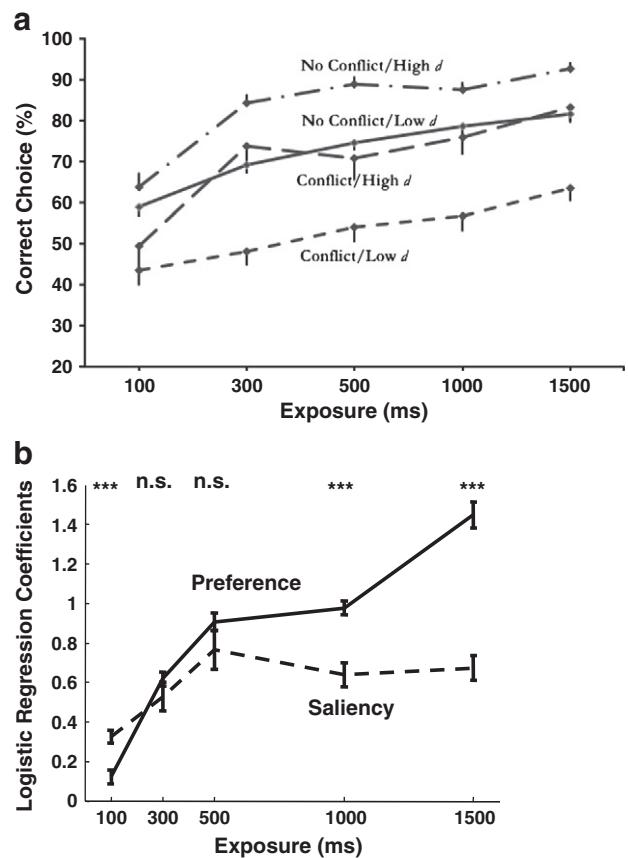


Fig. 6. External validity for visual saliency bias under cognitive load. (a) The percent of correct choices as a function of exposure, strength of preference, and whether or not there was a conflict between the preference and visual saliency. Error bars show S.E.M. (b) Estimated logistic regression coefficients for relative preference and visual saliency as a function of exposure time. Error bars represent S.E.M. (** $p<.001$).

of preference, i.e. choice difficulty ($F=13.21$, $p<.00001$), and conflict ($F=204$, $p<.00001$).

Fig. 6b depicts the estimated logistic regression coefficients for saliency and preference as a function of exposure time. The analysis shows that preference did not become the dominant factor influencing choice until 1000 ms ($p<.00001$, two-tailed t-test). As before, the influence of visual saliency on choice remained significant at all exposure durations ($p<.00002$; two-tailed t-test).

Finally, we once again found a strong effect of preference on choices under conditions of high brand dominance, i.e., strong preference (Fig. 7a). Still, it is worth noting that even then a significant fraction of choices favored the visually more prominent items, even though this was inconsistent with participants' preferences. At the same time, the visual saliency bias gained strength under conditions of low brand dominance, i.e., weak preference (Fig. 7b).

Discussion

The results of this final experiment show that the *visual saliency bias* is still significant and sizable when choices are

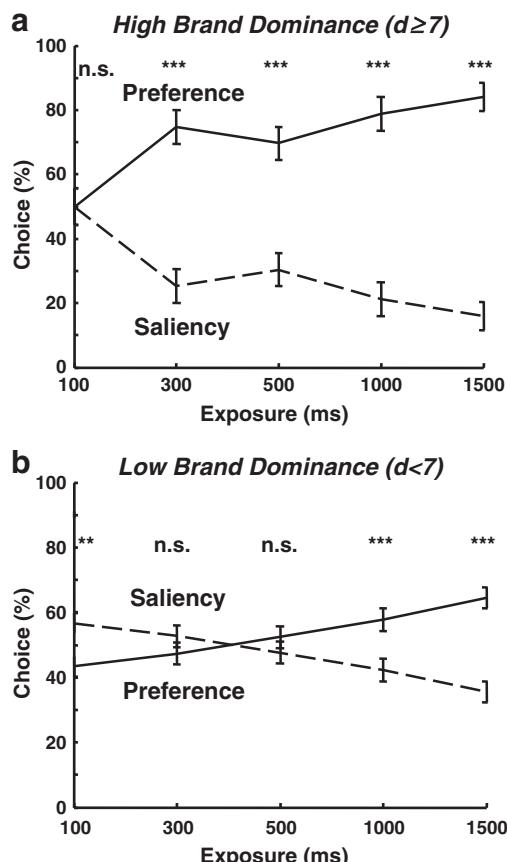


Fig. 7. Percent of choices on conflict trials (where one item is preferred but the other is more salient) for (a) easy choices with strong preferences for one of the items ($d \geq 7$) and (b) difficult choices with weak preference ($d < 7$). Error bars represent S.E.M. (***) $p<.001$.

indicated with button-presses, thus providing external validity for the results from Experiments 1 and 2.

General discussion

These results provide evidence for the existence of a sizable visual saliency bias, especially under the conditions of rapid decision making and cognitive load that characterize everyday decisions, such as many supermarket purchases. It is important to emphasize that our results are not due to the use of unusually fast decision speeds. For example, recent studies of choice using similar food stimuli found that participants typically make these types of decisions in 500–2000 ms (Krajbich et al., 2010; Hare, Camerer, & Rangel, 2009; Litt, Plassmann, Shiv, & Rangel, 2009). Another study found that participants could make accurate choices in less than 3 seconds in displays that included 16 items (Reutskaja, Nagel, Camerer, & Rangel, 2010). Note that these response times are measured from the time of stimulus presentation, whereas the response times reported above only include the time from stimuli offset to response initiation. Thus, the decision times in this paper are commensurate with those previously reported.

The results have implications for our understanding of marketing and consumer choice. First, the results identify a novel neurally-plausible mechanism through which marketing practices, such as the color of packaging or how the store shelf is lit, could have a sizable impact on individual decisions, even when those practices are not correlated with the consumers' preferences. Although the magnitude of the visual saliency bias amounts to less than a 1-point increase on the liking ranking scale at the longest exposures, this impact on choices can translate to large profits in competitive marketplaces where the profit margins are tight and participants can choose among many highly liked options. Second, the results advance our understanding of the mechanisms through which task-irrelevant cues, such as the relative visual saliency of objects, can affect choices (for other channels see Kotler, 1973; North, Hargreaves, & McKendrick, 1997; Dijksterhuis, Smith, van Baaren, & Wigboldus, 2005).

An important direction for future work is to understand which concrete visual features trigger visual saliency biases. For example, would it be possible to trigger this type of biases by introducing a low-brightness, mild-color package in a domain in which all competing products have red bright packages? Fortunately, our current understanding of the visual system allows us to extrapolate from our data to address questions like this. It is well known that the most prominent visual features (e.g., colors, brightness) in a given scene (e.g., a supermarket shelf) determine the visual saliency of items (e.g., bags of chips). This is the case because neurons at the retina, superior colliculus, lateral geniculate nucleus and early visual cortical areas are particularly tuned to such simple visual features. Visual saliency then causes certain items to "pop out" of a visual scene, leading to automatic attention toward these items (Koch & Ullman, 1985; Itti & Koch, 2001; Mannan et al., 2009). This automatic, bottom-up attention works based on the center-surround principle: what matters for attention is

feature contrast rather than absolute feature strength, e.g., a very bright item set against dimmer surroundings, or vice versa a dark item set against a bright background, will “pop out”. This suggests that what matters is to be visually different from the local surroundings, which induces an interesting problem of strategic competition in package design among competing brands.

Acknowledgments

We would like to thank Dr. Hilke Plassmann for useful comments on the earlier draft of this manuscript. Financial support from the Tamagawa Global Center of Excellence (MM), the Mathers Foundation (CK), the Gimbel Fund (CK), and the Moore Foundation (AR) is gratefully acknowledged.

References

- Armel, C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgement and Decision Making*, 3(5), 396–403.
- Armel, C., & Rangel, A. (2008). The impact of computation time and experience on decision values. *American Economic Review*, 98(2), 163–168.
- Bettman, J. R. (1979). *An information processing theory of consumer choice*. Reading, MA: AddisonWesley.
- Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, 34(4), 613–617.
- Dijksterhuis, A., Smith, P. K., van Baaren, R. B., & Wigboldus, D. H. J. (2005). The unconscious consumer: Effects of environment on consumer behavior. *Journal of Consumer Psychology*, 15, 193–202.
- Drolet, A., Luce, M. F., & Simonson, I. (2009). When does choice reveal preferences? Moderators of heuristic versus goal-based choice. *Journal of Consumer Research*, 36(1), 137–147.
- Fischer, B., & Weber, H. (1993). Express saccades and visual attention. *The Behavioral and Brain Sciences*, 16, 553–610.
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.
- Hare, T., Camerer, C., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vMPFC valuation system. *Science*, 324, 646–648.
- Henderson, J. M., Weeks, P. A., Jr., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology. Human Perception and Performance*, 25(1), 210–228.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, 2(3), 194–203.
- Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge: Cambridge University Press.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neuronal circuitry. *Human Neurobiology*, 4, 219–227.
- Kotler, P. (1973). Atmospherics as a marketing tool. *Journal of Retailing*, 49, 48–64.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in goal-directed choice. *Nature Neuroscience*, 13(10), 1292–1298.
- Krajbich, I., & Rangel, A. (2011). A multi-alternative drift diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108(33), 13852–13857.
- Litt, A., Plassmann, H., Shiv, B., & Rangel, A. (2009). Dissociating goal value and attention signals during simple decision making. *Cerebral Cortex*, 21, 95–102.
- Mannan, S. K., Kennard, C., & Husain, M. (2009). The role of visual salience in directing eye movements in visual object agnosia. *Current Biology*, 19(6), R247–R248.
- North, A. C., Hargreaves, D. J., & McKendrick, J. (1997). In-store music affects product choice. *Nature*, 390, 132.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42(1), 107–123.
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7), 545–556.
- Reutskaja, E., Nagel, R. M., Camerer, C., & Rangel, A. (2010). Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Economic Review*, 101, 900–906.
- Ratcliff, R., & Smith, P. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20, 873–922.
- Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, 6(12), 1317–1322.



First attention then intention

Insights from computational neuroscience of vision

Milica Milosavljevic and Moran Cerf
California Institute of Technology

Attention is a critical construct for anyone involved in marketing. However, research on attention is currently lacking in the marketing discipline. This is perhaps due to inherent difficulties in measuring attention. The current paper accentuates the importance of better understanding attention, and suggests studying attention as a two-component construct consisting of equally important bottom-up and top-down processes. While research on top-down attention has recently been undertaken by Pieters and Wedel (2004; 2007), the current paper introduces the field of computational neuroscience and its research on visual attention as a useful framework for studying bottom-up attention.

Attention is a prerequisite for all marketing efforts.

(Sacharin 2000)

Introduction

Research that integrates findings from cognitive psychology, cognitive neuroscience and marketing is in its infancy. Nevertheless, a few marketing researchers have ventured into this brave new world, which is expected to hold much potential for advertising research (Vakratsas & Ambler 1999). Merging cognitive neuroscience with research on consumer behaviour offers tremendous potential for growth in knowledge. This is especially so because ‘a great mismatch exists between the way consumers experience and think about their world and the methods marketers use to collect this information’ (Zaltman 2003).

Conveniently, neuroscience uses new technologies that make it possible to measure neurophysiological activity in order to study complex human behaviors (for a good overview of neuroimaging methods in terms

of relevance to consumer behavior research see Egidi *et al.* 2008; Plassmann *et al.* 2007). These tools have the potential to override the methodological problems of the former approaches (Plassmann *et al.* 2007), although whether this will actually occur remains to be seen. The hope is that these physiological measures will be used to augment, not replace, traditional research methods in order for marketing researchers to begin to more adequately validate and refute some of the long-debated theories of consumer behavior, and, by default, human behavior in general.

Within this broad research framework, of interest in terms of the present paper is one currently ignored, but perhaps crucial, aspect of consumer behaviour: selective attention. A recent review of how neuroscience can inform advertising (Plassmann *et al.* 2007) omits the construct of attention entirely. This is surprising given that marketing researchers have declared attention to be a prerequisite for all marketing efforts (see the introductory quote above, from Sacharin 2000). Thus, selective attention should be included within the neuroscience–marketing research framework.

Another omission in the emerging field that combines neuroscience and marketing is that insights from theoretical and computational neuroscience have yet to be introduced. Computational neuroscience combines what is known about the brain from neuroscience with the computing power available to simulate neuronal and psychological processes on a computer (Sejnowski *et al.* 1988). The goal of computational neuroscience is to develop algorithms that can simulate on a computer how the brain functions when we perform tasks (Smith & Kosslyn 2007). Although many computational models of memory, attention, learning and decision making have been introduced, the present paper focuses on perhaps the most developed: computational models of visual attention.

Thus, the current paper has two objectives. The first is to bring into the spotlight the construct of attention. The second is to introduce computational neuroscience of visual attention to the marketing field, and discuss its utility for understanding deployment of attention in the advertising context.

The construct of attention in an advertising context

Due to the plethora of communication channels, consumers are faced with an overabundance of information, where a typical consumer is exposed to

FIRST ATTENTION THEN INTENTION

several hundred (even several thousand) marketing messages daily. Not all of this information can be processed because of the limited capacity of the brain – known as the attentional bottleneck. In recognition of this cluttered environment, some researchers have declared that we are living in the attention economy, with attention being a scarce resource (Davenport & Beck 2002).

Not surprisingly, everyone involved with marketing knows the importance of getting consumers' attention. In advertising, the importance of attention is evidenced by its prominent position in many advertising models. Originating in 1898, the first formal advertising model, AIDA (Attention → Interest → Desire → Action), positioned attention as the first step that people go through when exposed to advertising and before making a purchase (Vakratsas & Ambler 1999). Furthermore, most hierarchy of effects models suggest that attention is a necessary step before higher-level processes.

Since the importance of attention to marketers is factual, it is surprising that marketing studies of attention are rare (Rosbergen *et al.* 1997). The little space that attention does receive is devoted to describing how it is measured, while little emphasis is placed on any conceptual discussion, as is described next.

The concept of attention

Even today, most marketing researchers' understanding of attention is similar to William James' view dating back to 1890: 'Every one knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought.' A single exception is the recent work of Pieters and Wedel (2004) and Rosbergen *et al.* (1997), who suggest that attention is a much more complex phenomenon than is currently studied in marketing.

Pieters and Wedel (2004) introduced two determinants (found in psychology and neuroscience) of attention to advertising. The two determinants are: (1) bottom-up and (2) top-down attention.

Bottom-up attention is a rapid, automatic form of selective attention that depends on the intrinsic properties of the input, such as its colour or intensity (Koch 2004). It is also known as saliency-based attention, indicating that the more salient an object, the higher the probability of it

being noticed. Top-down attention is a volitional, focal, task-dependent mechanism, often compared to a spotlight, that enhances processing of the selected item (Koch 2004).

In the present work, bottom-up processes are referred to as pre-attention, while top-down processes are referred to as focal attention. Thus, attention is viewed here as a two-step process, consisting of pre-attention and focal attention, although this is not necessarily a sequential process as top-down attention can sometimes moderate the bottom-up processes (Cerf *et al.* 2008).

Pieters and Wedel (2007) made important first steps towards improving our understanding of how top-down factors (i.e. consumers' goals – 'memorise the ads', 'collect brand information', 'evaluate product', etc.) may influence attentional deployment within magazine ads. However, it is also important to spark research on the effects of bottom-up, automatic attention. This type of research is virtually non-existent in the marketing literature. As the work of Pieters and Wedel is focused on print advertising, the same context will be used in the remainder of the current paper to demonstrate the importance of research on bottom-up attention.

The research on bottom-up attention may be especially important since, as with the other types of media, clutter is an imposing problem for magazine advertising. For example, a 318-page issue of *Glamour* magazine contains 195 pages of advertisements and 123 pages of editorial content (Clow & Baack 2004). Faced with such a large amount of clutter, consumers often have a singular goal: to avoid advertising. Even Pieters and Wedel (2007, p. 224) highlight that 'processing goals may have a lower likelihood of surfacing during the few seconds that consumers typically spend on ads during self-paced exposure', and that 'competitive clutter may favour reflexive control and hinder systematic goal control'. Also, they state that attention to ad objects is very low during free viewing of magazines because object salience, or the bottom-up driven attention, primarily determines attention during free viewing (Janiszewski 1998; Pieters & Wedel 2007). Thus, on many – although not all – occasions, bottom-up attention may be as close as advertisers can get to consumers before the top-down goal of 'ignore advertising' kicks in.

An earlier study by Pieters and Wedel (2004) serves well to further emphasise this point. They instructed more than 3,600 consumers to freely browse through magazines, and used eye tracking to measure where

FIRST ATTENTION THEN INTENTION

they directed their gaze. The experimental magazines included 1,363 print ads. They found that, on average, 95.7% of participants fixated at least once on ads, but that the lowest-scoring ad was skipped by 39% of the participants. Further, people spent on average 1.73 seconds with each ad, ranging from 0.037 seconds to 5.30 seconds.

This study shows that, even when people are seated in a laboratory and asked to look at a magazine, the time spent with ads is very low. It can be assumed that time spent with ads and the number of ads that people look at are even lower during natural magazine browsing when people's attention is also consumed by other factors in their environment.

Thus, the construct of attention in advertising should be studied based on the two-component framework, consisting of bottom-up and top-down attention. The biggest challenge in such undertaking is that while marketers recently began using improved measures of focal attention, measuring pre-attention is still challenging.

Measurement of attention

Perhaps the most common method of measuring attention to advertising is by using self-reported memory measures (e.g. 'To what extent did you pay attention to this ad?'). However, memory measures are poor indicators of what consumers pay attention to (Rosbergen *et al.* 1997), for at least two reasons. First, attention is known to precede awareness. Thus, it is possible that a stimulus was attended to, but has not reached the awareness stage, thus making it impossible for individuals to have it in their memory or to report it. Second, even if a stimulus was attended to, people are known to forget most of the stimuli they process.

A somewhat improved, although less frequently used, method of measuring attention in marketing is eye-tracking, where eye movements are recorded to indicate individuals' attentional patterns. The main weakness of eye-tracking, as currently used in advertising, is that findings are 'rather superficial' (Rayner *et al.* 2001, p. 220). For example, the size of the advertisement has been found to influence participants' looking times, which was pointed out by psychophysicists over a century ago (Tatler *et al.* 2005).

Thus, although some progress has been made in studying and measuring top-down attention, these methods do not account for bottom-up attention. What may prove to give life to research on bottom-up attention

is a branch of neuroscience known as computational neuroscience. To initiate this stream of research within marketing, computational neuroscience of visual attention is introduced next.

Computational neuroscience of visual attention

The goal of computational neuroscience is to relate the data of the nervous system to algorithms used by the brain to conduct higher-level human behaviours such as attention, learning, memory, emotions and decision making. The idea is to create a computer simulation of real human behaviour inspired by biological systems in the brain.

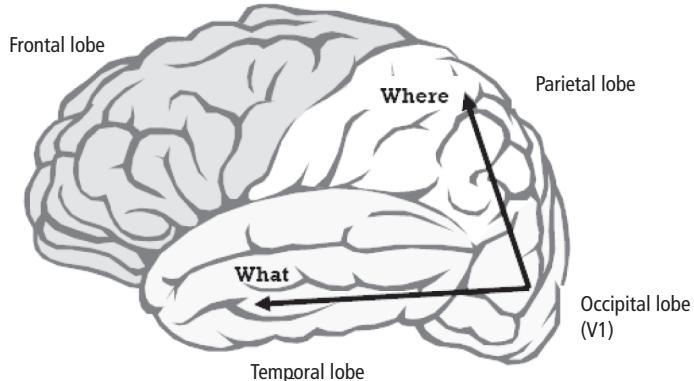
Computational brain modelling attempts to produce either (1) realistic brain models or (2) simplified compact brain models. A realistic brain model is a large-scale simulation that goes to the level of a single cell (Sejnowski *et al.* 1988). Since the model becomes much more realistic at the cellular level, it becomes less helpful in understanding its function at the nervous system level. Further, realistic simulation is computation-intensive, meaning that it requires a substantial computer power.

On the other hand, simplifying brain models are networks of brain cells or ‘neural networks’, which capture important principles of the functionality of a system (Sejnowski *et al.* 1988). Most importantly, neural networks are being used as models for psychological phenomena such as attention, emotions and decision making. Of those, perhaps the most realistic and advanced computational models are those that simulate visual attention.

Vision means ‘finding out what is where’ (Smith & Kosslyn 2007), and computational modelling attempts to provide algorithms that successfully locate and identify informative objects in a visual scene. Thanks to advances in neuroscience, it is known that certain brain areas give rise to visual attention, as shown in Figure 1.

Specifically, two separate cortical routes are involved in vision, giving rise to two streams of visual information (for a detailed review, see Koch 2004). Spatial deployment of attention (‘where’) is known as dorsal pathway. It proceeds from the primary visual cortex (V1) in the occipital lobe, through the posterior parietal cortex, and to the dorsolateral prefrontal cortex. Object recognition (‘what’) happens via the ventral pathway, which involves V1, the inferotemporal cortex, and the ventrolateral prefrontal cortex.

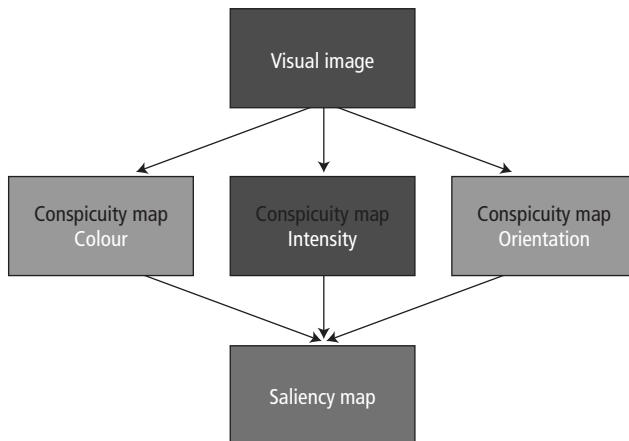
FIRST ATTENTION THEN INTENTION

Figure 1: Visual processing pathways

As mentioned earlier, attention is not given to all visual input. Since our visual environment is cluttered, attention serves as a processing bottleneck, allowing only a selected part of sensory input to reach visual awareness. This process depends on the two previously mentioned mechanisms: bottom-up and top-down attention. This two-component framework of visual attention was introduced by Treisman and Gelade (1980), and formed the basis for the development of computational models of visual attention. Guided by the idea that a visual scene is initially analysed automatically, based on the physical properties of objects in the scene, the first neurally plausible computational algorithm of bottom-up attention was developed by Koch and Ullman (1985), and later extended and implemented by Itti *et al.* (1998). The model is briefly introduced next.

Itti, Koch and Niebur's model of bottom-up attention and saliency

The model's flow (a rough outline of the model's architecture is depicted in Figure 2) begins by analysing the physical characteristics of objects in a given visual image. It analyses colour, intensity and orientation of objects, and sorts these into three conspicuity maps, which are grey-scale maps where brighter areas represent more salient locations while darker ones are less salient. One conspicuity map is created for each of the three characteristics: colour, intensity and orientation. For example, an object may be identified as highly salient because of its colour (the object's location

Figure 2: Computational model architecture

would be represented as a brighter area on the colour conspicuity map), while another object may be deemed salient due to its intensity (the objects' location would be brighter on the intensity conspicuity map).

The values from the three conspicuity maps are summed up into a saliency map, which is a two-dimensional topographic map that represents the saliency at every location in the visual image (Itti *et al.* 1998). The most salient locations are potential targets for visual attention (Schall & Thompson 1999). The most salient location is identified first. Then, this location is inhibited in a biologically motivated fashion and the next most salient location is determined, and so on. In this manner, attentional scan-paths are created for a given visual image (Itti 2004).

The model of bottom-up attention and saliency discussed here is a neurally based model – that is, it mimics human performance in a manner that is inspired by biological circuitry. Currently, it is probably the most widely used model of bottom-up visual attention (Cerf *et al.* 2007). It has been validated for the past decade on a number of classical visual search experiments, and was found to be ‘consistent with observations in humans’ (Duchowski 2002, p. 161). Recently, the model has been tested and improved in a number of contexts, including the presence of (1) motion, (2) faces and (3) text. These are discussed briefly next.

FIRST ATTENTION THEN INTENTION

First, while the saliency model was initially implemented on static images, it could easily be scaled to motion – taking frame-by-frame video data and analysing them as a single image. For example, recent additions to the model include flicker channels that allow for attention allocation to rapidly changing content that was shown to attract human fixations, and various semantic channels (Cerf *et al.* 2008).

Second, it was recently shown (Cerf *et al.* 2007) that adding a face channel – taken from existing face detector algorithms – can significantly enhance the predictions of the model when it comes to telling what human observers are looking at in an image. Also, and relevant to marketing research, when observers are looking at faces in an image this can also increase the amount of time spent viewing the image.

Finally, since we are exposed to text very often in our lives, it was shown in recent work by Cerf *et al.* (2008) that inclusion of a text channel in the saliency model further increases the ability of the model to predict what people are looking at. The performance of the model has been compared to the eye-tracking data collected from people exposed to a number of different images, and was found to be comparable.

It is important to note that the model is consistently found to perform comparable to results obtained from eye-tracking. The cost-benefit properties of such findings are the following: no expensive eye-tracking equipment is necessary, participants are not needed since the model simulates universal human bottom-up attention allocation, and significant time efficiency can be achieved as the algorithm can provide real-time analysis.

As this brief review of the current research on bottom-up attention shows, there has been much progress in modelling this process. It is important to remember that this sensory input is sometimes modulated by top-down, person and task-dependent input (Cerf *et al.* 2008). Computational models that include top-down cues are currently being developed by several research groups but are not as well developed as the models of bottom-up attention (Torralba & Oliva 2007; Cerf *et al.* 2008).

The following section discusses how computational modelling of visual attention may benefit both marketing theory and practice.

Potential contributions of computational neuroscience for marketing research

Using computational modelling of bottom up visual attention in marketing studies has the potential to make significant (1) theoretical, (2) empirical and (3) substantive contributions.

First, the conceptual understanding of attention will be enhanced by investigating factors that determine its key component: preattention. A better understanding of how to manipulate preattention will enable researchers to study its consequences, such as attitudes, intentions and/or choices. The effects of preattentive processing on attitudes towards the ad and brand were mentioned in the context of mere exposure effects when Janiszewski (1993, p. 376) highlighted the importance of investigating whether 'preattentive processes [are] instrumental in the formation of affective responses and, if so, how these preattentive processes operate'. About half a dozen marketing studies investigated the role of preattention within the mere exposure phenomenon, but no clear conclusions have been reached (Janiszewski 1993; Shapiro *et al.* 1997, 1999).

Second, although focal attention is being measured by somewhat improved methods of eye tracking, the measurement of preattention is still challenging. Yoo (2005) highlights this by stating that one of the most important emerging issues in the study of preattention to advertising is 'empirically detecting the existence of preattentive processing'. This can finally be addressed by introducing computational modelling of bottom-up attention, which *a priori* identifies objects that are likely to be preattentively processed by a viewer.

Finally, potential applications of computational modelling in the domain of advertising pretesting and evaluation are abundant. As an example, the following section demonstrates the utility of computational modelling of visual attention in the context of print advertising.

Application of computational modelling of visual attention to evaluation of print advertisements

As already suggested, many people purposely avoid looking at ads, or look at ads only very briefly. In such an environment, advertisers should be (and perhaps already are) trying to ensure that the key information in the ads is at least automatically, preattentively processed by consumers. This

FIRST ATTENTION THEN INTENTION

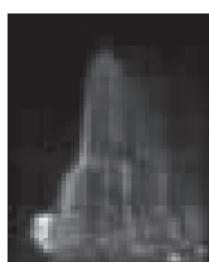
makes sense since it is known in psychology and neuroscience that automatic, preattentive processing is very rapid, occurring within less than one second of exposure to a visual scene (Quian Quiroga *et al.* 2008). The previously described computational model of visual attention (Itti *et al.* 1998) offers a tool that can be used in the design of print ads to ensure that key elements of an ad are salient, and thus more likely to be at least preattentively processed by viewers during these brief exposures.

To illustrate this process, two magazine ads from Procter & Gamble's Tide campaign, which won the 2007 Clio Award (available at www.clioawards.com/winners), are evaluated using a numerical computing software program – MATLAB – and the saliency algorithm of Itti *et al.* (1998), available at <http://ilab.usc.edu>.

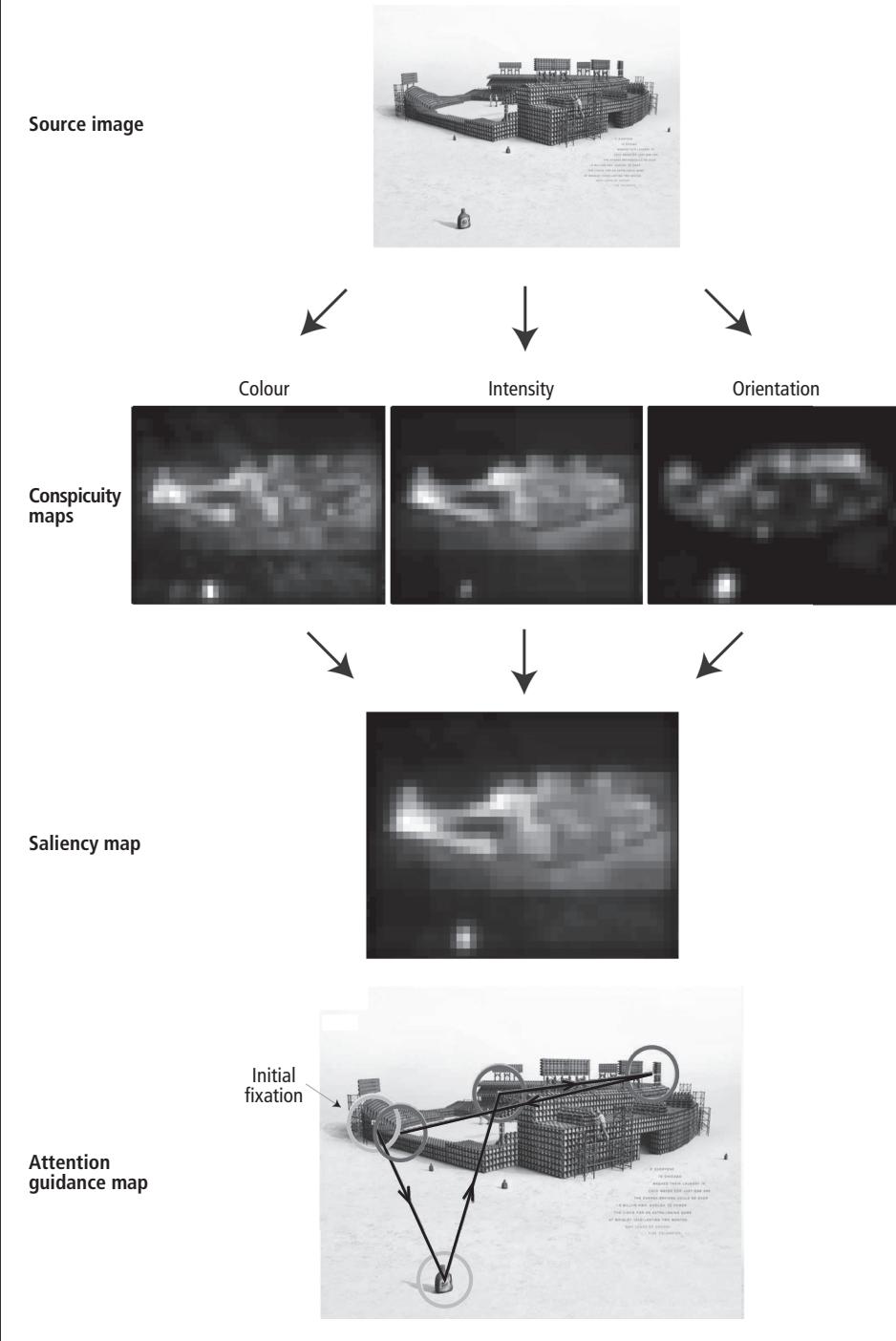
Figure 3 shows the computational modelling output for the first ad (original images are in colour, but are shown here in greyscale). The image of the ad is used as a visual input and is decomposed into three conspicuity maps (one for each of colour, intensity and orientation), which are then summed into a saliency map. The saliency map shows conspicuous ad objects, where the brighter the location the more noticeable the object. Based on the saliency map, and as shown in the 'Attentional Scanpath' image, the most salient locations (light-tinted circles) and the order in which attention shifts (black lines) are identified. Also, the time required for each shift of attention is calculated by the programme. In this example, the analysis simulates what an individual would preattentively process during the first half second of exposure to the ad.

The analysis shows that the most salient location is the left, dark side of the building (high intensity, shown on the intensity conspicuity map), which provides no information about the product or the brand. Perhaps a viewer may spend up to half a second on this shaded area before turning the page, without receiving any useful communication about the brand or the product. Since many people are unlikely to consciously pay attention to the ad, and computational modelling shows that the ad is not likely to be preattentively processed either, there is no reason to expect any positive advertising effects in this example.

The second ad comes from the same campaign and the output of computational modelling of this ad is shown in Figure 4. Once again, colour, intensity and orientation of objects in the ad are presented in the conspicuity maps, which are then summed into the saliency map.

Figure 3: Bottom-up attention to ad 1**Source image****Colour****Intensity****Orientation****Conspicuity maps****Saliency map****Attention guidance map**

FIRST ATTENTION THEN INTENTION

Figure 4: Bottom-up attention to ad 2

In this case, the product packaging at the bottom of the image (light-tinted circle) is the second most salient object (after the corner of the stadium), and even the third most salient location is within the product bottle (the label is salient due to its colour and then the top of the label becomes salient because of its orientation, as shown in Figure 4). Thus, even if a viewer consciously ignores the ad and spends as little as half a second on it before flipping the page, the brand name and product bottle are likely to be at least preattentively processed, opening a possibility for positive advertising effects. One can even argue that, although much research is needed to prove this hypothesis, a more relevant top-down goal – ‘I am out of and need to buy laundry detergent’ – may at this point override the ‘avoid advertising’ goal, thus resulting in focal attention to the ad.

Once again, it is important to note that a major advantage of using the computational model of bottom-up attention described here, if additional studies demonstrate its validity for advertising research, is that it does not require recruiting participants or employing time- and cost-intensive eye-tracking methodology.

Conclusions

The purpose of the current paper was twofold: (1) to highlight the importance of studying attention within the emerging research paradigm that combines marketing and neuroscience, and (2) to introduce the field of computational neuroscience to the marketing discipline.

First, the paper points out that attention currently receives virtually no space in the neuroscience–marketing literature, even though attention is a necessary step for all other marketing efforts. More so, it was emphasised that studying attention as a two-component construct consisting of a combination of bottom-up and top-down processes provides a useful framework for increasing understanding of the ways by which attention operates. In such a quest, the context of consumer behaviour proves to be a very relevant and natural way to study and better understand attentional processes.

Second, using computational modelling of visual attention to simulate early attention on a computer offers much potential for improving conceptual understanding and methods of measuring preattention, as well as a host of opportunities for application in the field.

FIRST ATTENTION THEN INTENTION

In summary, the current paper identified a gap in the marketing and, more specifically, advertising literature, and thus has opened numerous research possibilities, some of which are discussed briefly next.

Directions for future research

Future research should uncover how preattention operates; for example, which physical characteristics of objects result in preattention, and when are they effective? Once a better understanding of preattention has been achieved, the studies that assess the relationship between preattention and focal attention, as well as between preattention and attitudes, emotions and decision-making processes, should follow. This will enhance our understanding of the concept of attention, its antecedents and its consequences.

The present paper has demonstrated the utility of the computational modelling of visual attention in a specific advertising context – magazine advertising. Some recent preliminary studies that are yet to receive attention of marketing researchers are mentioned next.

First, in the work of Torralba and Oliva (2007), it was shown that people can quickly identify the 'gist' of the image which can bias attention allocation. For example, attention is allocated faster to an image of pedestrians (expected to be walking on the ground) when they are shown at the bottom of the image than to an image of pedestrians located at the top of the image (Torralba & Oliva 2007). In the context of magazine advertising, for example, if we want to emphasise the shoes of the model in an ad, they should probably be placed in the bottom margin of the page, thus increasing the probability that the object will be noticed. Future research in advertising context is needed to test to which extent attention allocation depends on such general properties of a scene.

Second, the role of attention allocation by bottom-up-driven saliency models has been studied in the context of video gaming. Peters and Itti (2007) recently showed that some bottom-up-driven attention mechanisms may not only govern ad viewing, but even computer game playing. Future research in this context is warranted given current escalation of in-game advertising.

Finally, in a preliminary study, the computational model of bottom-up attention and saliency (Itti *et al.* 1998) was used to design banner ads on a

website in order to make them more or less salient (Milosavljevic 2007). In an experiment where all other factors were controlled and only the saliency of the banner ad was manipulated, consumers' attitudes towards the banner ad were progressively enhanced as they spent increasingly more time on the website in the condition where the banner ad was designed to be salient, while no such change was observed for non-salient banners (Milosavljevic 2007).

It is important to note that, as argued earlier, recall and recognition rates for the target ad were very low (less than 20%), even in this extreme condition where most people spent two to three minutes exposed to the ad on various web pages. This provides further support for the argument that people often purposely avoid looking at the ads, and once again points to the importance of studying the confluence of bottom-up and top-down attention. The hope is that the current paper will motivate both marketing academicians and practitioners to better understand the construct of attention and to perhaps do so by utilising knowledge and tools from cognitive and computational neuroscience.

References

- Cerf, M., Frady, P. & Koch, C. (2008) Using semantic content as cues for better scanpath prediction. *Eye Tracking Research & Applications*, pp. 143–146.
- Cerf, M., Harel, J., Einhaeuser, W. & Koch, C. (2007) Predicting human gaze using low-level saliency combined with face detection. *Advances in Neural Information Processing Systems*, **120**, pp. 241–248.
- Clio Awards (2007) www.clioawards.com/winners.
- Clow, K.E. & Baack, D. (2004) *Integrated Advertising, Promotion, and Marketing Communications*. Upper Saddle River, NJ: Prentice Hall.
- Davenport, T.H. & Beck, J.C. (2002) *The Attention Economy: Understanding the Currency of Business*. Boston, MA: Harvard Business School Press.
- Duchowski, A.T. (2002) *Eye Tracking Methodology: Theory and Practice*. London: Springer-Verlag.
- Egidi, G., Nusbaim, H.C. & Cacioppo, J.T. (2008) Neuroeconomics: foundational issues and consumer relevance, in Haugtvedt, C.P., Herr, P.M. & Kardes, F.R. (eds) *Handbook of Consumer Psychology*. New York: Taylor & Francis Group, pp. 1177–1207.
- Itti, L. (2004) Modeling primate visual attention, in Feng, J. (ed.) *Computational Neuroscience: A Comprehensive Approach*. Boca Raton FL: Chapman & Hall/CRC, pp. 607–619.

FIRST ATTENTION THEN INTENTION

- Itti, L., Koch, C. & Niebur, E. (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**(11), pp. 1254–1259.
- James, W. (1890) *The Principles of Psychology*. New York: Dover Publications.
- Janiszewski, C. (1993) Preattentive mere exposure effects. *Journal of Consumer Research*, **20**(3), pp. 376–392.
- Janiszewski, C. (1998) The influence of display characteristics on visual exploratory search behavior. *Journal of Consumer Research*, **25**(3), pp. 290–301.
- Koch, C. (2004) *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts & Company Publishers.
- Koch, C. & Ullman, S. (1985) Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, **4**(4), pp. 219–227.
- Milosavljevic, M. (2007) *Antecedents and Consequences of Preattention*. Working Paper, California Institute of Technology.
- Peters, R. & Itti, L. (2007) Congruence between model and human attention reveals unique signatures of critical visual events. *Advances in Neural Information Processing Systems*, **20**, pp. 1145–1152.
- Pieters, R. & Wedel, M. (2004) Attention capture and transfer in advertising: brand, pictorial, and text-size effects. *Journal of Marketing*, **68** (April), pp. 36–50.
- Pieters, R. & Wedel, M. (2007) The goal control of attention to advertising: the Yarbus implication. *Journal of Consumer Research*, **34**(2), pp. 224–233.
- Plassmann, H., Ambler, T., Braeutigam, S. & Kenning, P. (2007) What can advertisers learn from neuroscience? *International Journal of Advertising*, **26**(2), pp. 151–175.
- Quiroga, R., Mukamel, R., Isham, E.A., Malach, R. & Fried, I. (2008) Human single-neuron responses at the threshold of conscious recognition. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(9), pp. 3599–3604.
- Rayner, K., Rotello, C.M., Stuart, A.J., Keir, J. & Duffy, S.A. (2001) Integrating text and pictorial information: eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, **7**(3), pp. 219–226.
- Rosbergen, E., Pieters, R. & Wedel, M. (1997) Visual attention to advertising: a segment-level analysis. *Journal of Consumer Research*, **24**(3), pp. 305–314.
- Sacharin, K. (2000) *Attention! How to Interrupt, Yell, Whisper & Touch Customers*. New York: John Wiley & Sons, Inc.
- Schall, J.D. & Thompson, K.G. (1999) Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience*, **22**, pp. 241–259.
- Sejnowski, T.J., Koch, C. & Churchland, P.S. (1988) Computational neuroscience. *Science*, **241**(4871), pp. 1299–1306.
- Shapiro, S., MacInnis, D.J. & Heckler, S.E. (1997) The effects of incidental ad exposure on the formation of consideration sets. *Journal of Consumer Research*, **24**(1), pp. 94–104.
- Shapiro, S., MacInnis, D.J., Heckler, S.E. & Perez, A.M. (1999) An experimental method for studying unconscious perception in a marketing context. *Psychology and Marketing*, **16**(6), pp. 459–477.
- Smith, E.E. & Kosslyn, S.M. (2007) *Cognitive Psychology: Mind and Brain*. Upper Saddle River, NJ: Prentice Hall.

- Tatler, B.W., Baddeley, R.J. & Gilchrist, I.D. (2005) Visual correlates of fixation selection: effects of scale and time. *Vision Research*, **45**(5), pp. 643–659.
- Torralba, A. & Oliva, A. (2007) The role of context in object recognition. *Trends in Cognitive Sciences*, **11**(12), pp. 520–522.
- Treisman, A.M. & Gelade, G. (1980) A feature-integration theory of attention. *Cognitive Psychology*, **12**(1), pp. 97–136.
- Vakratsas, D. & Ambler, T. (1999) How advertising works: what do we really know? *Journal of Marketing*, **63**(1), pp. 26–43.
- Yoo, C.Y. (2005) Preattentive processing of web advertising. Doctoral dissertation, University of Texas, Austin, TX.
- Zaltman, G. (2003) *How Customers Think: Essential Insights into the Mind of the Market*. Boston, MA: Harvard Business School Press.

About the authors

Milica Milosavljevic is associate researcher at the California Institute of Technology. Her research interests are at the intersection of psychology of consumer behavior and cognitive/computational neuroscience. Her recent research focuses on the interplay of visual processing/visual attention and decision-making. She received her PhD from Florida Atlantic University and has taught several marketing courses, including consumer behaviour, at both undergraduate and graduate level.

Moran Cerf is a researcher at the California Institute of Technology pursuing a PhD in Neuroscience. His research focuses on the ways by which the brain encodes information of high level sensory and cognitive processes at the neuronal level, and the interplay between those processes and conscious attention allocation. He studies attention using direct recordings from human brain as well as eye-tracking in order to develop models of attention allocation in the visual system.

Address correspondence to: Milica Milosavljevic, Division of the Humanities and Social Sciences, California Institute of Technology, MC 228-77, 1200 E. California Blvd, Pasadena, CA, 91125, US.

Email: mmilosav@hss.caltech.edu

Brand preference affects the threshold for perceptual awareness

THOMAS ZOËGA RAMSØY^{1,2*} and MARTIN SKOV^{1,2}

¹Center for Decision Neuroscience, Department of Marketing, Copenhagen Business School, Solbjerg, Plads 3, 2000 Frederiksberg, Denmark

²Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Kettegård Allé 30, 2650 Hvidovre, Denmark

ABSTRACT

Despite decades of scientific scrutiny, much is still unknown about the effects that brands have on perception. Brands are known to lead to changes in attention and mnemonic processing and by altering emotional preferences they imbue products with value. Less, however, is known about the exact mechanism through which this occurs. Here, a novel and unexpected finding is provided in which subjective brand preference alters the likelihood that a brand name will be consciously seen. By presenting brand names at brief durations, and having them respond using a graded evaluation of conscious perception, the Perception Awareness Scale, it is found that brand names for which there is either a positive and negative preference, subjects report seeing the name more clearly. Interestingly, and much to the contrary of studies of basic emotions, this effect is strongest for positive preference. Our results are discussed in light of other studies in consumer psychology and consciousness science. Copyright © 2013 John Wiley & Sons, Ltd.

The basic ambition of using a brand is to distinguish a product or service positively from other competing products. Thus, through branding efforts, brand managers hope to get consumers to identify the product with a specific identity, to influence the consumers' ability to recall and recognize the product (among other products), and to associate the product with qualities that are positively perceived (Wänke *et al.*, 2007; Plassmann *et al.*, 2012). To accomplish these goals, brands must engage a larger number of our mental systems, including perception and awareness, memory and cognition, and emotion.

Exactly how all these complex systems come together to make brands work their magic remains obscure. One important aspect of brand psychology that has recently started to come under scrutiny is the study of the mechanisms of why brands are able to modulate the subjective value of an object. It is well known that brands hold the ability to influence how the consumer values a product. For instance, consumers report that the same physical drink tastes better if they are told it is a specific beer brand than if they drink it without this knowledge (Allison and Uhl, 1964), and they display a willingness to pay significantly higher prices for original labels than for identical knock-offs. In other words, a brand can enhance a consumer's preference for, and perhaps even hedonic experience of (Ariely and Norton, 2009; Ariely and Berns, 2010), a product *beyond* its physical properties. Even with exact copies, being associated with a brand can make a consumer value one object over the other item that is not associated with the brand. Thus, brands "infuse" value into the product, as it were.

In studying these mental effects, neuroimaging studies have gone some way to identify the brain regions involved in processing brand value. For example, in a seminal paper, McClure *et al.* (2004) showed that when subjects drink an

identical liquid and were told on some trials that they were drinking Coca Cola and other trials that they were drinking Pepsi Cola, the impression that they were drinking Coca Cola led the subjects not only to rate the taste of the liquid as more pleasant but also produced enhanced neural activity in regions involved in memory and attention (hippocampus and dorsolateral prefrontal cortex). When not receiving brand information, subjective ratings of taste pleasure were positively related to activation in the orbitofrontal cortex (OFC). Similarly, in a functional Magnetic Resonance Imaging study Kirk *et al.* (2009) led subjects to believe that abstract paintings were either from a prestigious art museum in Denmark, or not. Again, a positive bias in rating for paintings associated with the "art gallery" paintings corresponded to stronger brain activity in the OFC, as compared with paintings not associated with such brand information.

Yet, knowing which anatomical regions of the brain that underlie brand valuation does not in itself explain *why* a brand may modulate the perceived value of an object. Presumably, learned and memorized associations of the brand elicit positive or negative predictions of how pleasant the object will be, something that has been shown possible to occur unconsciously (Pessiglione *et al.*, 2008) influencing the neural processes involved in computing the actual experienced value of the object (Plassmann *et al.*, 2008; Santos *et al.*, 2011).

In terms of consumer responses to brands, at least two processes are at play. On the one hand, overt experiences of brand value are reflected in verbal reports and are reflected in the specific engagement of the OFC, and as such denote what we here call the "feelings" accompanying the perception of a brand (Plassmann *et al.*, 2012). On the other hand, as with abstract symbols, brands may lead to automatic behavioral and physiological responses that occur without accompanying awareness, yet nevertheless may affect subsequent processing of brand-related information, even in the absence of awareness (Pessiglione *et al.*, 2008), which we here call "emotions." What lies at the very borderline between unconscious emotions and conscious feelings is the subject of this paper.

*Correspondence to: Thomas Z. Ramsøy, Decision Neuroscience Research Group, Department of Marketing, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark.
E-mail: tzr.marktg@cbs.dk

Preferences and visual perception

Knowing that brands carry value, that this value can infuse the product the brand gets associated with, and that a specific “hedonic” brain region – the OFC (Rolls, 2000; O’Doherty *et al.*, 2001; Plassmann *et al.*, 2010; Grabenhorst and Rolls, 2011) – plays an important role in this process, which raises a number of interesting new questions. Specifically, since recent studies have demonstrated an interconnection between affective processing and consciousness (e.g. Thomsen *et al.*, 2011) and that brands have been shown to be able to modulate affective processing in the brain, one may wonder if brand value also holds the ability to affect how a product is consciously perceived. The argument would thus be the following: (1) At the individual level, a brand becomes associated with a particular value, a learning that can both have conscious and unconscious elements. (2) Seeing a physical representation of the brand, be it the brand name, a brand logo, or some other brand representation, evokes this value by influencing the brain’s affective system. Notably, even unconscious presentation of the brand can evoke such responses. (3) The evoked affective value modulates the perceptual processing of whatever object the brand is associated with, including visual mechanisms involved in making a percept conscious.

From studies on the effects of basic emotions, the subjective valence of an object is known to affect both perceptual processing and behavioral responses (Öhman and Mineka, 2001). Compared with neutral stimuli, more attentional resources are allocated to the perceptual computation of emotional, especially negative, objects (Algom *et al.*, 2004; Estes and Verges, 2008). The result is an increase in detection accuracy (Nasrallah *et al.*, 2009). Furthermore, neuroimaging research has demonstrated that this difference in how emotionally relevant stimuli are processed, is associated with a difference in neural activation of structures involved in perceptual processing (Lang *et al.*, 1998; Lane *et al.*, 1999; Mourão-Miranda *et al.*, 2003).

Reward expectancy and other emotions, in turn, are known to regulate attention. Studies in psychology and cognitive neuroscience have demonstrated alterations in attention to visual objects following operant and classical conditioning (Platt and Glimcher, 1999; Glimcher, 2003; Dorris and Glimcher, 2004; Sugrue *et al.*, 2004). Recently, Serences (2008) demonstrated that, in the absence of changes in overt eye movements, stimuli with a reward history produced stronger engagement of the visual systems of the brain, as indexed by functional Magnetic Resonance Imaging. Indeed, one plausible and empirically supported view is that regions involved in emotion and saliency, such as the amygdala, can affect sensory processing through top-down attentional modulation (e.g. Vuilleumier *et al.*, 2004).

Why would there be this interaction between subjective value, affect, and perceptual processing? The effect of stimulus valence on perceptual processing is probably rooted in the evolutionary benefit of having sensory information processing guided by biological relevance (Vuilleumier, 2005). Although the exact anatomical and functional relation between the neural processes involved in visual perception

and emotion remains unclear (Pessoa and Adolphs, 2010), perceptual factors associated with biological relevance such as saliency or predictability have been shown to directly influence activity in brain regions such as the amygdala, and correlate with avoidance-prone behavior (Herry *et al.*, 2007), just as amygdala activity induced by fearful faces has been shown to modulate neural activity in early visual cortex (Vuilleumier *et al.*, 2004).

Advantages of conscious processing

As the review of the relationship between emotional processing and perception shows, it is biologically possible that brand preference could influence how brands themselves or objects associated with the brand are consciously perceived. But what, if true, would the implication of this be?

Consciousness research during the past several decades has demonstrated several significant cognitive differences between conscious and unconscious processing. One of the main findings emerging from this research is that unconscious processes exhibit a limited and rigid information processing capacity, make use of specialized processors, and exert little effect on overall thinking and action (Baars, 2002; Baars *et al.*, 2003; Shanahan and Baars, 2005). By contrast, conscious processes are associated with a large and flexible information processing capacity, make use of dynamic processors, and exert a much larger effect on thinking and behavior (Baars, 2002). In other words, whether a stimulus is consciously seen or not has a tremendous influence on the effects it can have on thinking and behavior. Although there is robust evidence showing that subliminal perception – stimuli processed by the sensory system but not consciously experienced – can influence thoughts, feelings and actions (Wänke *et al.*, 2007; Chartrand *et al.*, 2008) – consciously perceived information enables a much stronger effect on flexible and creative behaviors (Shanahan and Baars, 2005).

As a consequence, stimuli that reach consciousness are more likely to affect behaviors relevant to consumer decisions, including memory, preference, and decision making. In the informationally crowded environment of today’s product and media landscapes, only a fraction of the stimuli being processed by the perceptual system gains full access to consciousness and the benefits of such process. Hence, understanding the factors and mechanisms responsible for whether or not a given stimulus obtains access to consciousness is of great value for our understanding of consumer psychology, as well as providing a better insight into the relationship between preference and awareness.

Here, evidence is presented in which the likelihood that a stimulus becomes conscious depends upon how much it is subjectively preferred. Using an influential and validated measure of consciousness, the Perception Awareness Scale (Ramsøy and Overgaard, 2004; Christensen *et al.*, 2006; Overgaard *et al.*, 2006; Sandberg *et al.*, 2010), we find that that the threshold for conscious processing of a stimulus is affected by the individual level of preference for the stimulus being briefly displayed.

Hypotheses

The current knowledge about the effects of preference on consciousness stems from work performed on the stimuli with a direct emotional value and is the only source of prior knowledge from which one can base experimental hypotheses. On the basis of the aforementioned studies (Öhman and Mineka, 2001; Dijksterhuis and Aarts, 2003), two hypotheses were proposed.

Hypothesis 1: Brand names for which the subject have either a positive or a negative preference have a higher likelihood of being consciously perceived, compared with brands for which the subject has a neutral preference.

This hypothesis stresses the general idea that emotional stimuli, through their regulatory effect on attention, are more likely to gain access to conscious processing. Here, this insight is extended by testing whether stimuli for which a subject has a specific positive or negative preference, will also produce a preconscious attention response that leads to an increased likelihood for conscious experience of the stimulus. Notably, in the study by Dijksterhuis and Aarts (2003), there was no neutral condition to test whether positive stimuli were more easily perceived than neutral ones. Hence, a neutral condition – brand names for which the subject is indifferent – was employed to test the main hypothesis.

This leads to the second hypothesis. On the one hand, the recent report that negative stimuli were more easily detected than stimuli words (Dijksterhuis and Aarts, 2003) leads us to expect that a similar effect is present for brand preference:

Hypothesis 2: Brands for which subjects have a negative preference (dislike) are more likely to be consciously seen than brands that the subject has a positive preference (like).

However, two facts make an alternate hypothesis equally plausible. First, studies have suggested that positive preference may indeed be related to a stronger effect on visual attention than negative preference. For example, Serences (2008) reported a positive bias in the effect of reward history on neural activation related to visual attention. More specifically, there was a significant effect of selected (i.e. positive preference) compared with unselected (i.e. negative preference) items on such visual activation. Second, using preferences such as brands may rely on different processing mechanisms not comparable with basic emotional cues such as snakes, spiders, or facial expressions. Hence, the alternative hypothesis that should be considered as equally probable is that there is no difference between positive and negative preferences, or even that positive preference exerts a stronger effect than negative preference, on conscious detection.

METHOD

To this end, 49 subjects (age mean/std = 25.7/2.7, 26 women, and 46 right handed) were recruited from the region of Copenhagen, Denmark. Subjects signed an informed consent and were positioned in front of a computer monitor (the distance was approximately 60 cm; the screen type was a

cathode ray tube screen running at 75 Hz to ensure brief stimulus presentation). First, each subject underwent a brand rating test in which they were sequentially shown brand names on-screen and were asked to rate their brand preference using a 7-point scale (1 = strongly dislike, 7 = strongly like), and a second 7-point scale rating for their knowledge of the brand (1 = do not know, 7 = know very well). In all, each subject rated 104 brand names by pressing the numeric buttons on a keyboard in a self-paced manner. The brand names were selected on the basis of two criteria: first, they should be among the most prominent brands available in Denmark (taken from a national list of the most prominent Danish brands); and second, the brand names should be approximately at the same length (4–9 characters) to avoid systematic differences in visual masking of these stimuli. To maximize statistical power and reduce individual variance in scoring, the two most positive (6–7) and most negative (1–2) liking scores were collapsed into single categorical scores of positive and negative brand liking, respectively, and used these scores in the statistical analyses. A score of 4 was used as the neutral brand preference.

Subjects were instructed that in the following test, they would be shown different brand names at different durations and that their task was to report the perceived clarity of the brand name. Subjects were also told that there would be null events, but that their main task was to report the clarity of their perception by adhering to the Perception Awareness Scale (PAS) (Ramsøy and Overgaard, 2004). All brand names were presented multiple times and in pseudorandom order (randomized but avoiding the presentation of the same brand name two or three times in a row) for each subject. Each brand name was presented for approximately 13, 52, and 91 ms in a forward- and backward masking paradigm. Subjects reported the clarity of their visual experience of each name by using three of the possible responses of the PAS: having “no experience,” a “vague experience,” or a “clear experience” of the stimulus (Figure 1).

To analyze the effect of brand preference on conscious experience (PAS score), we ran an Ordinal Logistic Regression, using PAS score as the dependent variable, and with individual brand rating (positive, neutral, and negative) as an independent variable, and with brand knowledge and stimulus duration as additional regressors. Reaction times below 100 ms and above 5000 ms were excluded. Statistical analysis was performed in JMP version 9.0 (SAS Institute Inc.)

First, the main effect of brand preference rating on PAS scores was analyzed, with stimulus duration and knowledge scores as covariates. Second, the analysis was divided into separate duration times. In a post-hoc analysis, a more complex model was tested in which the interaction between duration and preference on PAS score was included.

RESULTS

The statistical analysis showed that the whole model was highly significant ($\chi^2 = 9558.6$, $R^2 = 0.316$, $p < 0.0001$). Looking at the individual effects, we find highly significant effect of brand preference on PAS scores ($\chi^2 = 8.5$,

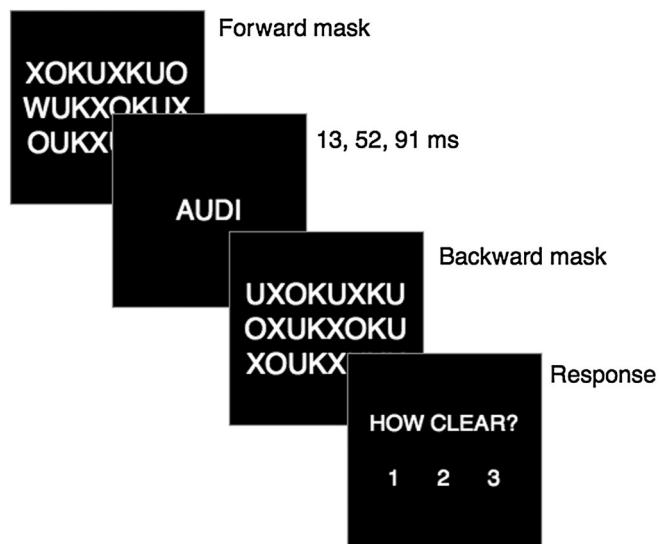


Figure 1. The experimental procedure. Subjects first saw a forward mask, followed by a brief presentation of the brand name for either 13, 52 or 91 ms. Following a backward mask, subjects were asked to rate the clarity of their experience using the Perception Awareness Scale.

$\beta = -0.036$, $p = 0.0036$). Additional effects were found for knowledge ($\chi^2 = 20.6$, $\beta = -0.054$, $p < 0.001$) and stimulus duration ($\chi^2 = 9549.6$, $\beta = -0.076$, $p < 0.0001$). As Figure 2A illustrates, the PAS effect was due to higher scores for positive and negative brand preference scores relative to neutral preference. A post-hoc analysis looking at nonlinear effects (2nd order polynomial) of preference on PAS judgements was found to provide a slightly better model ($\chi^2 = 13.5$, $\beta = -0.01$, $p = 0.0003$); thus, reflecting higher PAS scores for low and high preferences, relative to neutral preference.

To further probe this effect, the difference between positive, negative, and neutral individual brand preferences on PAS score was analyzed using direct comparisons. Here, the analysis demonstrated a significant difference between negative compared with neutral brand ratings ($\chi^2 = 31.1$, $p < 0.001$) and an even stronger effect for positive compared with neutral brand ratings ($\chi^2 = 129.5$, $p < 0.0001$). By directly comparing the effects of positive and negative brand ratings, it was found that positive brand ratings were associated with significantly higher PAS scores than negative brand ratings ($\chi^2 = 19.5$, $p = 0.013$).

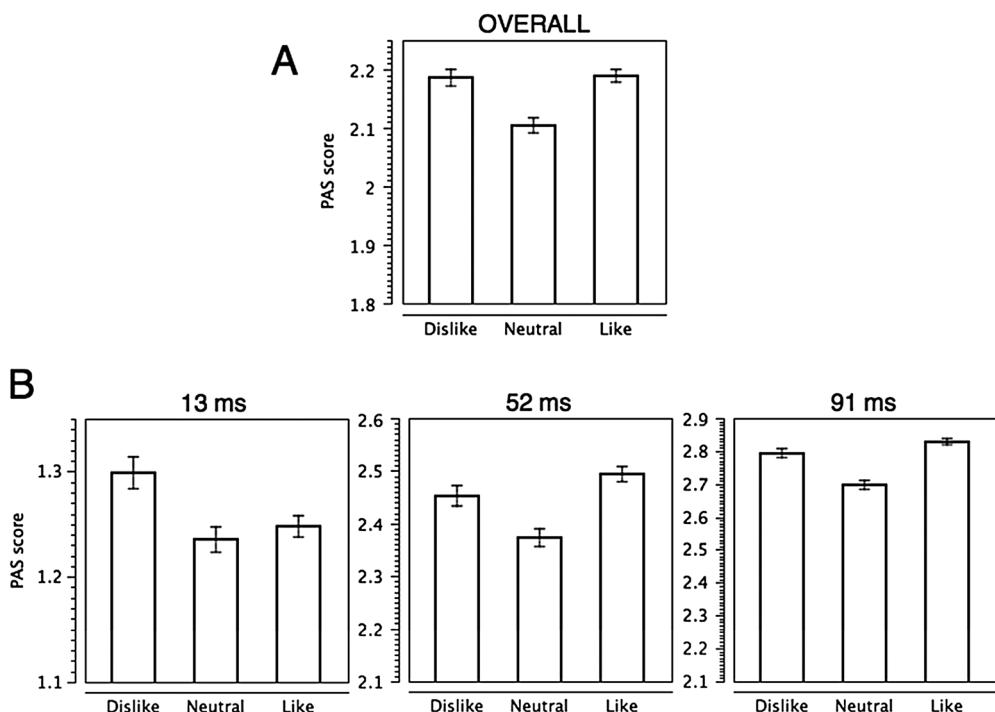


Figure 2. Effects of brand preference on Perception Awareness Scale (PAS) scores. The main effect (A) demonstrates that stimuli associated with either positive or negative preference are generally associated with higher PAS scores than brand names for which subjects are indifferent. Looking at these effects for each stimulus duration (B), the effects are most pronounced for 91 ms duration and not significant for the 13 ms condition. Bar charts showing mean value, whiskers denote confidence interval.

A follow-up analysis was performed to look at the interaction between preference and stimulus duration on PAS scores, still modeling the main effects of preference and duration individually, and with knowledge as covariate. Here, we find that the model is highly significant ($\chi^2=9608.3$, $R^2=0.317$, $p < 0.0001$) and that in addition to the effects seen for duration, preference and knowledge, there is a preference duration interaction ($\chi^2=49.6$, $\beta < -0.01$, $p < 0.0001$). As seen in Figure 2B, at 13 ms negative preference was associated with higher PAS scores, whereas at 52 and 91 ms both positive and negative preferences were associated with higher PAS scores.

Collapsing the 52 and 91 ms presentations, direct comparisons between brand preferences demonstrated that both positive and negative brand preferences were related to significant higher PAS scores (positive: $\chi^2=88.5$, $p < 0.0001$; negative: $\chi^2=85.4$, $p < 0.0001$) when compared with neutral preference. Again, positive brand preference showed a significantly higher mean PAS score than negative preference ($\chi^2=88.5$, $p < 0.0001$).

General discussion

To our knowledge, this is the first study to show an effect of individual preference on the clarity of conscious perception. By using a forward- and backward masking paradigm with variable stimulus duration, and employing a sensitive measure to capture graded conscious perception, the results demonstrate a strong effect of individual brand preference on the detection and identification of briefly presented brand names. Notably, these results are distinct from the effects of brand knowledge and stimulus duration. In particular, brands for which subjects had either a positive or negative preference were more likely to be consciously perceived, compared with brands for which subjects had a more neutral relationship. This finding provides support for Hypothesis 1, and suggests that processes associated with computing the hedonic value of a stimulus influence the neural mechanisms mediating attention and ultimately conscious perception. Our results lend support to the notion that emotional processing plays an important role in the consciousness (Dijksterhuis and Aarts, 2003) and extends this notion by demonstrating that the effects hitherto found only for basic emotions can also be produced when using more complex and acquired emotional processes (i.e., subjective preferences).

Second, counter to our original a priori hypothesis, the data show that negative emotions are not more likely to provide access to consciousness than positive emotions. On the contrary, brands with a high positive preference (liking) were significantly more likely to be perceived consciously than brands with a low preference rating (disliking). Thus, Hypothesis 2 must be rejected. This result is at odds with previous studies using basic emotions (e.g., Öhman and Mineka, 2001; Dijksterhuis and Aarts, 2003; Gaillard *et al.*, 2006) and warrants further examination and replication. As noted, this inconsistency may be related to at least two specific factors. First, the nature of the stimuli used in the present study is different when compared with other studies. Although earlier studies have employed stimuli associated with direct reward and punishment, the use of brand names

for which there is individual and acquired preference represents a different level of processing. In other words, brands may engage the brain's valuation system differently than basic emotions and therefore produce different effects on attention and consciousness. Indeed, it may be possible that the effects seen here are more related to other psychological phenomena, such as the mere exposure effect (e.g., Janiszewski, 1993). This idea should be examined further in future studies. Second, and related to this, the recent study by Serences (2008) in which subjects underwent a value-based decision-making paradigm, it was found that stimuli with a positive reward history produced a stronger response in the visual system than stimuli with a negative reward history. As brand preference can indeed be seen as the result of an ongoing learning of brand-reward contingency, the results provided by Serences (2008) may indeed provide a plausible neurocognitive mechanism for which brand preference affects visual attention and, through this, conscious access. However, to fully embrace this hypothesis, more studies are needed.

Our results provide novel insights into the relationship between emotions and consciousness in consumer psychology. Studies of the brain bases of consciousness have often treated the threshold to consciousness as a stable variable and assumed that the threshold is unaffected by cognitive or emotional factors. Using the duration of stimulus presentation as the key variable has been dominant. Most studies have employed a cutoff at around 50 ms at which subjects start to report perceiving the stimulus (Ramsøy and Overgaard, 2004; Christensen *et al.*, 2006; Overgaard *et al.*, 2006; Kouider and Dehaene, 2007). Notably, theoretical accounts have suggested that consciousness is an all-or-none phenomenon caused by a neural bifurcation in which information changes from being processed locally (e.g. in the primary visual areas) to engaging a widespread brain network including thalamic, frontal, and parietal regions (Baars *et al.*, 2003; Christensen *et al.*, 2006). This view of consciousness has been used to suggest that information from one processor (e.g. vision) is temporarily given influence on a widespread network, often referred to a "global workspace" of the brain (Baars, 2002; Baars *et al.*, 2003; Christensen *et al.*, 2006), ultimately leading to an improved emotional and cognitive processing and larger effect on behavior.

It is thought that such neural globalization effects work by influencing other processes related to behavior. Indeed, a percept that becomes conscious is more likely to be acted upon (e.g. through approach or avoidance behaviors) than non-perceived information. In other words, conscious processing is suggested to have a stronger power on mental processes and behavior than subliminal processes. Although the literature consistently demonstrates an effect of subliminal processes on behavior, by far the strongest influence on behavior has been found through overt, conscious processing (Baars, 2002; Baars *et al.*, 2003). Given this, the effect of emotion on the threshold to consciousness should also affect subsequent behavioral effects. As emotions work by affecting behavioral inclinations of withdrawal or approach behaviors, the added effect of subjective preference on the conscious

threshold could increase the likelihood that the stimulus is detected, and hence, has an increased influence on psychological processes and behavior. Future research should build on our findings and directly compare the psychological and behavioral effects of conscious and unconscious brand perception.

We chose to use brands written in standardized letters to allow easier forward- and backward masking of the brands. This follows the tradition of consciousness research and psychology in which it has been found that the masking of divergent stimuli is more problematic (Kouider and Dupoux, 2004; Kouider *et al.*, 2010). Moreover, our choice of simple text was to avoid additional differences between brands in color, fonts, and other variables that might explain the behavioral effects. Nevertheless, as recent studies have suggested that unconscious processing of abstract symbols can affect choice-related behaviors (Pessiglione *et al.*, 2008); this suggests that forthcoming studies should seek to test whether simplified symbolic representations of brands may have the same, or even bigger, effects on the relationship between brand preference and conscious detection.

Furthermore, it should be noted that our approach could not be performed with more complex and brand-specific stimuli such as brand logos. As brand logos have several differences in terms of brightness, contrast, and color composition, there are two main difficulties in testing the threshold to consciousness with this method. First, the forward- and backward mask method used relies on the ability to make a composite mask that can mask the stimuli. With words using the same fonts, this can be carried out reliably. However, following the vast literature on visual masking (see for example Kouider and Dupoux, 2004; Pessiglione *et al.*, 2008; Kouider *et al.*, 2010), this is not possible when using different complex images, as the mask will not work sufficiently for all stimuli. Second, and following the same point, because of the visual properties of the logos themselves (brightness, contrast, color composition, etc.), some stimuli may be less successfully masked. This may thus elevate the possibility that the experimental design will be suboptimal and invalid. Taken together, our results are limited by the use of a single aspect of brands, that is, their names, and future studies should use alternative means to study the effect of brand preference on the threshold to consciousness.

It should also be noted that our results provide independent insights into the nature of consciousness. Several studies have argued against the dichotomous view of consciousness and studies suggest that subjects are neither fully conscious nor fully unconscious of a stimulus (Ramsøy and Overgaard, 2004; Christensen *et al.*, 2006; Sandberg *et al.*, 2010). Rather, studies in neuropsychology and experimental psychology have clearly demonstrated cases in which subjects or patients report having sensory experiences that cannot be classified as either clearly conscious or unconscious. Instead, there are instances in which the observer reports having vague, glimpse-like experiences of a stimulus. Patients sometimes report this as “black on black,” or liken it to “a mouse under the carpet” (Zeki and Ffytche, 1998; Ffytche and Zeki, 2011). Given this, one can contend that conscious perception can be classified into an instance of conscious

detection, in which the observer reports seeing something but not being able to report on the exact identity of the item being shown, and conscious *identification*, in which the observer has a clear percept of what was shown (for a review see Ramsøy and Overgaard, 2004).

In terms of brand communication, our data provide novel suggestions for the importance of improving brand preference. In particular, besides the effects that overt brand liking can have on products and information, our data suggest that brand preference affects attentional mechanisms. In today’s increasingly crowded consumer communication space, such findings are of utmost importance. For example, the American Association for Advertising Agencies (Britt and Adams, 2007) provided a conservative estimate of a typical exposure to over 600 ads per day for an adult (although if you count all exposures whenever you pass a brand you might run up to 20,000 exposures per day). Of those advertisements, only about 10% were explicitly noticed and recalled. Our results suggest that which ads are recalled is not purely coincidental but that there is a significant contribution of individual preference – especially strong disliking and liking – upon the likelihood that a brand will be detected and as such be remembered. This suggests that from a brand communications perspective, inducing strong emotional responses in consumers has an added benefit that it increases the likelihood of that brand being detected in a crowded environment. Although we do not think that inducing strong brand disliking in consumers will have consumer benefits, there may be a space in which negative communication may have an impact on detection, such as in health-related communication.

Our study lends further support to the use of the PAS, a method for assessing the clarity of perception. The PAS was originally conceived in 2004 (Ramsøy and Overgaard, 2004) in which it was demonstrated that above chance levels of guessing the features of a stimulus was only possible when subjects reported having detected the stimulus, but chance levels of guessing when reporting having seen nothing. Employing neuroimaging measures, the PAS score was related to a graded neural response in regions known to be involved in conscious perception (Christensen *et al.*, 2006). Taken together, our results lend further support to the use of the PAS in studying consciousness and extends these results by demonstrating that emotions lead to a parallel shift in perceptual clarity, as assessed by the PAS. That is, for the brief stimulus duration, emotionally neutral percepts tend to be unseen by the subject, but stimuli that bears on a subjective preference are more likely to be reported as “glimpses.” Likewise, at middle stimulus durations, brand names with emotional salience are more likely to be reported as clear experiences compared with brand names for which the subject has a neutral preference.

CONCLUSION

The results presented here are, to our knowledge, the first to demonstrate that subjective hedonic value – measured as subjective preferences for brands – leads to a stronger

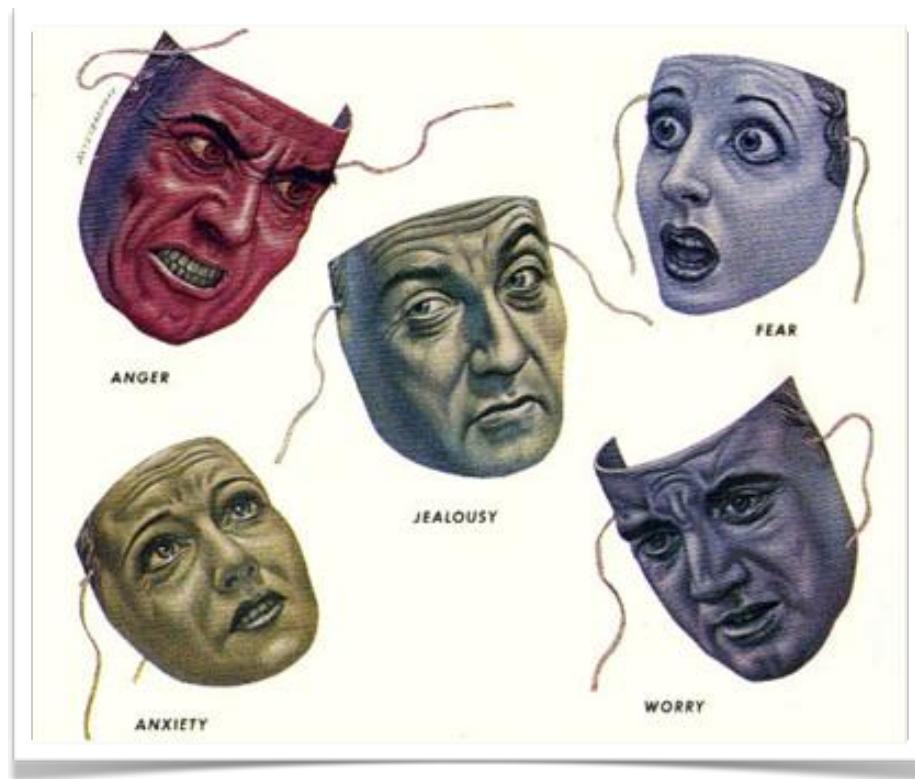
likelihood that information associated with the brand will be processed consciously. Furthermore, by extending the study to include both positive (like), negative (dislike), and neutral preferences (indifferent), previous research is extended and show that positive preference has a larger impact on consciousness than both negative and neutral preferences. These results both improve our understanding of the relationship between emotions, valence, attention and consciousness, and in particular, our understanding of how brands affect emotional processing, leading to significant changes in attention and subsequent cognitive processing.

REFERENCES

- Algom D, Chajut E, Lev S. 2004. A rational look at the emotional Stroop phenomenon: generic slowdown, not a Stroop effect. *Journal of Experimental Psychology. General* 133: 323–338.
- Allison RI, Uhl KP. 1964. Influence of beer brand identification on taste perception. *Journal of Marketing Research* 1(3): 36–39.
- Ariely D, Berns GS. 2010. Neuromarketing: the hope and hype of neuroimaging in business. *Nature reviews. Neuroscience* 11(4): 284–92.
- Ariely D, Norton MI. 2009. Conceptual consumption. *Annual Review of Psychology* 60: 475–99.
- Baars BJ. 2002. The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Sciences* 6(1): 47–52.
- Baars BJ, Ramsøy TZ, Laureys S. 2003. Brain, conscious experience and the observing self. *Trends in Neurosciences* 26(12): 671–675. doi:10.1016/j.tins.2003.09.015
- Britt SH, Adams SC. 2007. How many advertisements is a person exposed to in a day? American Association of Advertising Agencies. <https://ams.aaa.org/ewebs/upload/FAQs/adexposures.pdf>
- Chartrand TL, Huber J, Shiv B, Tanner RJ. 2008. Nonconscious goals and consumer choice. *Journal of Consumer Research* 35: 189–201.
- Christensen MS, Ramsøy TZ, Lund TE, Madsen KH, Rowe JB. 2006. An fMRI study of the neural correlates of graded visual perception. *NeuroImage* 31(4): 1711–25. doi:10.1016/j.neuroimage.2006.02.023
- Dijksterhuis A, Aarts H. 2003. On wildebeests and humans: the preferential detection of negative stimuli. *Psychological Science* 14(1): 14–18.
- Dorris MC, Glimcher PW. 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44: 365–378.
- Estes Z, Verguts M. 2008. Freeze or flee? Negative stimuli elicit selective responding. *Cognition* 108: 557–565.
- Ffytche DH, Zeki S. 2011. The primary visual cortex, and feedback to it, are not necessary for conscious vision. *Brain: A Journal of Neurology* 134(Pt 1): 247–57.
- Gaillard R, Del Cul A, Naccache L, Vinckier F, Cohen L, Dehaene S. 2006. Nonconscious semantic processing of emotional words modulates conscious access. *Proceedings of the National Academy of Sciences of the United States of America* 103: 7524–7529.
- Glimcher PW. 2003. The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience* 26: 133–179.
- Grabenhorst F, Rolls ET. 2011. Value, pleasure and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences*.
- Herry C, Bach DR, Esposito F, Di Salle F, Perrig WJ, Scheffler K, Lüthi A, Seifritz E. 2007. Processing of temporal unpredictability in human and animal amygdala. *Journal of Neuroscience* 27: 5958–5966.
- Janiszewski C. 1993. Preattentive mere exposure effects. *The Journal of Consumer Research* 20(3): 376–392.
- Kirk U, Skov M, Hulme O, Christensen MS, Zeki S. 2009. Modulation of aesthetic value by semantic context: an fMRI study. *NeuroImage* 44(3): 1125–32.
- Kouider S, Dehaene S. 2007. Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1481): 857–875.
- Kouider S, Dupoux E. 2004. Partial awareness creates the “illusion” of subliminal semantic priming. *Psychological Science* 15(2): 75–81. doi:10.1111/j.0963-7214.2004.01502001.x
- Kouider S, de Gardelle V, Sackur J, Dupoux E. 2010. How rich is consciousness? The partial awareness hypothesis. *Trends in Cognitive Sciences* 14(7): 301–7. doi:10.1016/j.tics.2010.04.006
- Lane RD, Chua PM, Dolan RJ. 1999. Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia* 37: 989–997.
- Lang PJ, Bradley MM, Fitzsimmons JR, Cuthbert BN, Scott JD, Moulder B, Nangia V. 1998. Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology* 35: 199–210.
- McClure SM, Li J, Tomlin D, Cypert KS, Montague LM, Montague PR. 2004. Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44(2): 379–87. doi:10.1016/j.neuron.2004.09.019
- Mourão-Miranda J, Volchan E, Moll J, de Oliveira-Souza R, Oliveira L, Bramati I, Gattass R, Pessoa L. 2003. Contributions of stimulus valence and arousal to visual activation during emotional perception. *NeuroImage* 20: 1955–1963.
- Nasrallah M, Carmel D, Lavie N. 2009. Murder, she wrote: enhanced sensitivity to negative word valence. *Emotion* 9: 609–618.
- O'Doherty J, Kringlebach ML, Rolls ET, Hornak J, Andrews C. 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience* 4(1): 95–102.
- Öhman A, Mineka S. 2001. Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review* 108: 483–522.
- Overgaard M, Rote J, Mouridsen K, Ramsøy TZ. 2006. Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition* 15(4): 700–8. doi:10.1016/j.concog.2006.04.002
- Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, Frith CD. 2008. Subliminal instrumental conditioning demonstrated in the human brain. *Neuron* 59(4): 561–7.
- Pessoa L, Adolphs R. 2010. Emotion processing and the amygdala: from a ‘low road’ to ‘many roads’ of evaluating biological significance. *Nature Reviews Neuroscience* 11: 773–782.
- Plassmann H, O'Doherty J, Shiv B, Rangel A. 2008. Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences of the United States of America* 105(3): 1050–4.
- Plassmann H, O'Doherty JP, Rangel A. 2010. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *The Journal of Neuroscience* 30(32): 10799–808.
- Plassmann H, Ramsøy TZ, Milosavljevic M. 2012. Branding the brain – a critical review and outlook. *Journal of Consumer Psychology* 22: 18–36.
- Platt ML, Glimcher PW. 1999. Neural correlates of decision variables in parietal cortex. *Nature* 400: 233–238.
- Ramsøy TZ, Overgaard M. 2004. Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences* 3: 1–23.
- Rolls ET. 2000. The orbitofrontal cortex and reward. *Cerebral Cortex* 10(3): 284–294.
- Sandberg K, Timmermans B, Overgaard M, Cleeremans A. 2010. Measuring consciousness: is one measure better than the other? *Consciousness and Cognition* 19(4): 1069–78. doi:10.1016/j.concog.2009.12.013
- Santos JP, Seixas D, Brandão S, Moutinho L. 2011. Investigating the role of the ventromedial prefrontal cortex in the assessment of brands. *Frontiers in Neuroscience* 5: 77.
- Serences JT. 2008. Value-based modulations in human visual cortex. *Neuron* 60: 1169–1181.

- Shanahan MP, Baars B. 2005. Applying global workspace theory to the frame problem. *Cognition* 98(2): 157–176.
- Sugrue LP, Corrado GS, Newsome WT. 2004. Matching behavior and the representation of value in the parietal cortex. *Science* 304: 1782–1787.
- Thomsen KR, Lou HC, Joensson M, Hyam JA, Holland P, Parsons CE, Young KS, et al. 2011. Impact of emotion on consciousness: positive stimuli enhance conscious reportability. *PLoS one* 6(4): e18686.
- Vuilleumier P. 2005. How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences* 9: 585–594.
- Vuilleumier P, Richardson MP, Armony JL, Driver J, Dolan RJ. 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience* 7: 1271–1278.
- Wänke M, Herrmann A, Schaffner D. 2007. Brand name influence on brand perception. *Psychology and Marketing* 24(1): 1–24.
- Zeki S, Ffytche DH. The Riddoch syndrome: insights into the neurobiology of conscious vision. *Brain* 1998; 121: 25–45.

EMOTIONS & FEELINGS



One of the important concepts in marketing and consumer insights is emotions, or feelings...? What is really the basic difference between emotions and feelings? Are they just two concepts for the same phenomenon, or do they reveal something more fundamental about how our minds are organized?

To make matters worse, emotions and feelings are used interchangably in academia and industry. Here, I will tease apart what the basic differences are, and serve some central articles pertaining to this distinction.

Put simple, **emotions** can be said to be neural and bodily responses to inner or outer events. Your startle response to a sudden noise or sweating in your palms when nervous are good examples of emotional responses. **Feelings**, on the other hand, should be reserved to the process of having a conscious experience of being in a particular emotional state. This distinction suggests that we can have many emotional responses without an accompanying feeling, but we cannot have feelings without an accompanying emotion.

ARTICLE

Received 12 Nov 2013 | Accepted 1 Jul 2014 | Published 29 Jul 2014

DOI: 10.1038/ncomms5567

OPEN

Audience preferences are predicted by temporal reliability of neural processing

Jacek P. Dmochowski^{1,†}, Matthew A. Bezdek², Brian P. Abelson³, John S. Johnson³, Eric H. Schumacher² & Lucas C. Parra¹

Naturalistic stimuli evoke highly reliable brain activity across viewers. Here we record neural activity from a group of naive individuals while viewing popular, previously-broadcast television content for which the broad audience response is characterized by social media activity and audience ratings. We find that the level of inter-subject correlation in the evoked encephalographic responses predicts the expressions of interest and preference among thousands. Surprisingly, ratings of the larger audience are predicted with greater accuracy than those of the individuals from whom the neural data is obtained. An additional functional magnetic resonance imaging study employing a separate sample of subjects shows that the level of neural reliability evoked by these stimuli covaries with the amount of blood-oxygenation-level-dependent (BOLD) activation in higher-order visual and auditory regions. Our findings suggest that stimuli which we judge favourably may be those to which our brains respond in a stereotypical manner shared by our peers.

¹ Department of Biomedical Engineering, City College of New York, 160 Convent Avenue, New York, New York 10027, USA. ² School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, Georgia 30332, USA. ³ Harmony Institute, 54 West 21st Street, New York, New York 10010, USA.
† Present address: Department of Psychology, Stanford University, 450 Serra Mall, Stanford, California 94305, USA. Correspondence and requests for materials should be addressed to J.P.D. (email: jdmochow@stanford.edu).

Predicting the behaviour of large groups is inherent to such diverse processes as forecasting election results, anticipating the reception to upcoming films, and foreseeing the effects of changes to laws or policies. Meanwhile, the prediction of individual behaviour is a pillar of neuroscience, with a recent focus on the study of naturally occurring behaviours. Previous investigations have identified the neural correlates of individual preferences^{1–5}, subjective values⁶ and choices^{7–9} by measuring the functional magnetic resonance imaging (fMRI)-derived blood-oxygenation-level-dependent (BOLD) signal in regions-of-interest while subjects perform experimental tasks. Here we ask whether the neural activity of multiple individuals may collectively predict the behaviour of large groups.

Previous works aimed at predicting population trends from brain activity have employed the amplitude of a neural signal, typically the BOLD, as a readout of future behaviour^{4,9,10}. Such an approach implicitly assumes that the strength of neural response in a fixed region correlates with behavioural measures. More recently, however, a growing link is emerging between the reliability of neural processing (that is, correlation across repeated presentations of the stimulus) and natural behaviours. Indeed, naturalistic audiovisual stimuli have been shown to elicit highly reliable neural activity across multiple viewers¹¹, with the level of such inter-subject correlation (ISC) linked to successful memory encoding¹² and effective communication between individuals¹³. ISC is increased during scenes marked by high arousal and negative emotional valence^{11,14}, and is strongest for familiar and naturalistic events¹⁵. In addition to these fMRI studies, recent work found that engaging narrative stimuli yield high levels of ISC in the evoked encephalographic responses of a small sample of viewers^{16,17}.

Given the evidence linking ISC— inherently a group measure—to brain states characterized by heightened affect, attention and memory retention, we suspected that the agreement in neural responses may serve as a suitable predictor for subsequent population behaviour. Specifically, we hypothesized that the level of neural reliability elicited by a naturalistic stimulus in a small sample would be predictive, to some degree, of behavioural responses reflecting engagement or interest of a large population.

Broadcasts of popular television shows or advertisements serve as a convenient framework for testing our hypothesis: in the social media age, the responses of large audiences are captured in online networks such as Twitter, Facebook and YouTube. We leverage this to explore the link between neural and behavioural responses. Namely, we recruited a sample of 12–16 naïve subjects and presented them with stimuli which had been previously aired and for which we compiled aggregated measures of the population response. We imaged brain activity during this exposure, employing electroencephalography (EEG) which captures broad patterns of activity on the time scale of neuronal processing, allowing us to measure reliability in short temporal segments. To further characterize the observed reliability, we subsequently performed an EEG-informed fMRI activation study to identify brain areas which are systematically more (or less) active during stimuli which elicit greater ISCs in the EEG. Most importantly, we found a statistically significant link between the neural reliability in the sample and preferences of large audiences within and across contemporary audiovisual stimuli. Our findings suggest that behavioural responses of large groups to natural stimuli may be robustly predicted from the reliability in corresponding neural responses of a small sample of individuals.

Results

We sought a stimulus eliciting time-varying and readily available viewer responses across a large population. To that end, we

considered the premiere broadcast of a popular television series ('The Walking Dead', AMC, 2010) in conjunction with two metrics which capture the audience's response to the original broadcast in a time-resolved manner.

An online service which collects Twitter traffic information was employed to obtain a comprehensive listing of time-stamped, stimulus-relevant tweets, which originated during the airing of the episode. Meanwhile, 16 study participants representative of the series' target demographic were recruited to view the episode while having their neural activity recorded with high-density EEG.

The stimulus was partitioned into its 190 constituent scenes (ranging in duration from 1.4 to 300.5 s, with a median length of 17 s), where a scene was defined as an aggregate of shots (that is, uninterrupted sequences of frames) comprising a distinct narrative event. For each scene, we computed the frequency of elicited tweets. To account for the non-negativity and heavy-tailed distribution of Twitter activity¹⁸, we logarithmically transformed the tweet rate to yield the time series shown in Fig. 1a, which defines our dependent measure.

Meanwhile, we sought to measure the amount of neural reliability evoked by each scene in our sample of participants. Rather than computing reliability in an electrode-to-electrode fashion, we first performed a dimensionality reduction technique which projects the neural responses from all subjects onto a space which maximizes the ISCs across our sample (see Methods for details of computation). When measured in this optimized space, the bulk of the reliability is captured in just a few dimensions (that is, 3). The resulting scene-by-scene neural reliability was then regressed onto our dependent measure, yielding the predicted log tweet frequency (see equation (4) in Methods) shown in Fig. 1b.

The neural reliability experienced by the sample throughout each scene explains 16% of the variance in audience log tweet frequency (Fig. 1c; $r = 0.40$, $P = 6.1 \times 10^{-7}$, $N = 190$, P -value computed using the analytic distribution of the sample correlation coefficient¹⁹, 95% confidence interval on r : (0.26,0.51) computed using the bootstrap²⁰). It is worthwhile to note that while tweeting is a delayed behavioural response, the observed neural reliability is driven by immediate short-term responses (reliability was calculated for activity 0.5 Hz or higher; see Methods and Fig. 4).

On the basis of previous findings suggesting an association between ISC and narrative quality, novelty and coherence^{11,16,21}, we suspected that neural reliability may also predict viewership size. To that end, we obtained minute-by-minute Nielsen ratings stemming from the original broadcast (including advertisements), resulting in a time series conveying audience size and defining our dependent measure (Fig. 2a).

The decision to continue viewing may depend, in part, on recent viewing history. As a result, we opted not to correlate reliability instantaneously with viewership. Instead, we formed our neural reliability measure using ISCs computed over the prior 3 min of viewing. We then regressed the resulting time series onto the minute-by-minute viewership, yielding the predicted time series shown in Fig. 2b. Neural reliability explains 36% of viewership (Fig. 2c; $r = 0.60$, $P = 7.1 \times 10^{-8}$, $N = 86$; 95% confidence interval on r : (0.45,0.71)).

There are two evident sources of variability in the Nielsen ratings: a sudden drop in ratings during advertisements, and a gradual decay due to declining audience retention. To determine if the measured correlation is driven by the obvious variation from intervening advertisements, we repeated the calculation but omitting the advertising segments. Reliability explains 34% of the variance during programming alone ($r = 0.58$, $P = 2.6 \times 10^{-5}$, $N = 62$, 95% confidence interval on r : (0.33,0.75)).

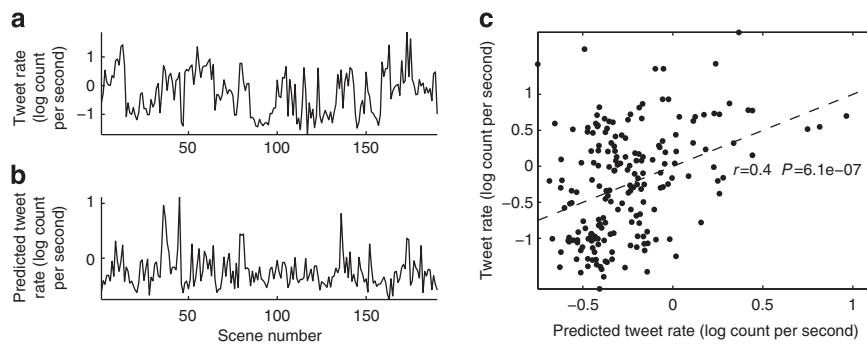


Figure 1 | Neural reliability predicts scene-by-scene tweet frequency. (a) Log frequency of tweets related to each scene ($N=190$) of a popular television broadcast. (b) Log tweet frequency as predicted from the scene-by-scene neural reliability measured across 16 participants during subsequent viewing of the episode in the laboratory (see equation (4) in Methods). (c) Neural reliability explains 16% of the variance in the log tweet rate ($r=0.4$, $P=6.1 \times 10^{-7}$, $N=190$; 95% confidence interval on r : (0.26,0.51)). Dashed line represents regression from predicted to actual log tweet rate.

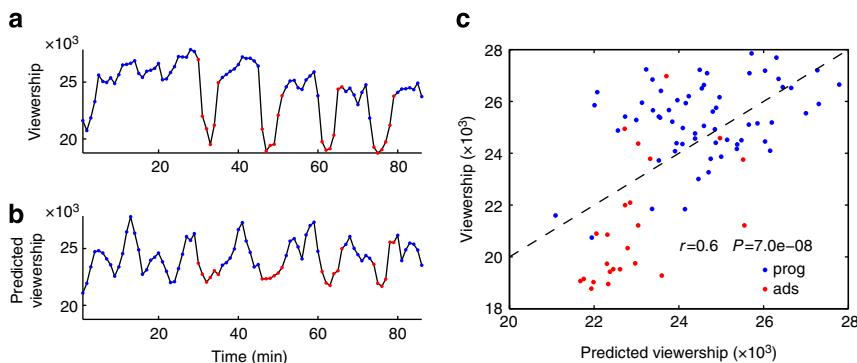


Figure 2 | Neural reliability predicts viewership size. (a) Viewership size during broadcast of television show as measured by Nielsen ratings. Programming (blue) is interrupted by advertising (red). (b) Viewership as predicted from the neural reliability exhibited by 16 participants viewing the same programming. (c) Neural reliability explains 36% of the variance in viewership size ($r=0.60$) when including both periods of programming (blue) and advertising (red), while accounting for 34% of the variance during programming alone ($r=0.58$). Dashed line denotes regression from predicted to actual viewership (including ads); prog ads, programming advertisements.

The gradual drop in viewership size observed here is typical of the free-viewing environment of the general audience (that is, being able to change the channel at any time). This contrasts with the laboratory environment in which participants are asked to view the entire episode. To compensate for this mismatch in viewing conditions, we removed the linear trend in the viewership size and found even stronger correlations (complete broadcast: $r=0.68$, $P=4.9 \times 10^{-11}$, 95% confidence interval on r : (0.56,0.77); programming only: $r=0.66$, $P=3.1 \times 10^{-7}$, 95% confidence interval on r : (0.45,0.82)). In other words, neural reliability explains 43% of the variance in viewership size during programming after accounting for the drop in retention.

We also considered the effect of the temporal window size (that is, 3 min) used to define reliability on the prediction accuracy. As shown in Supplementary Fig. 1, the strength of the relationship between neural reliability and viewership exhibits a broad peak at a window size of 3–4 min when predicting ratings during both programming and advertisements, while increasing monotonically from 1 to 6 min when excluding ads (see also Supplementary Note 1). In addition, the correlation of viewership size with neural reliability is insensitive to which of the two age categories provided by Nielsen is being predicted (Supplementary Table 1 and Supplementary Note 2).

Both the tweet frequencies and Nielsen ratings considered above quantify audience response during a single programme. Audience preferences are often expressed not within but across competing programming. We wanted to test the ability of the

sample neural reliability to predict across-stimuli preferences. We thus obtained the results of a popular online survey occurring annually, in which a large number of participants view and subsequently rate a series of advertisements (SuperBowl commercials). We randomly sampled 10 ads from the 2012 version of this survey and recruited a new set of $N=12$ volunteers to view these ads while recording their EEG. Subjects also provided their own preference rating following the recording. For each advertisement, we computed the neural reliability from the ISCs in the neural responses of the sample (see Methods for details). We found a strong and statistically significant correlation between neural reliability and the population ratings (Fig. 3a, circled markers; $r=0.90$, $P=9 \times 10^{-5}$, $N=10$, 95% confidence interval on r : (0.76,0.97)). Given this surprisingly strong correlation, we sought to validate the results on a new stimulus set, repeating the experiment with the 2013 series of ads while employing the same 12 participants. The neural reliability correlated significantly with the population ratings (Fig. 3a, triangle markers; $r=0.73$, $P=0.014$, $N=10$; 95% confidence interval on r : (-0.06,0.95); the drop in correlation from 2012 is driven by a single advertisement, see Supplementary Note 3). By combining all 20 advertisements viewed by each study participant, neural reliability explains 66% of the variance in population ratings (Fig. 3a; $r=0.81$, $P=3 \times 10^{-6}$, $N=20$, 95% confidence interval on r : (0.50,0.92)). Intriguingly, neural reliability explains just 26% of the sample's own preferences (Fig. 3b; $r=0.51$, $P=0.02$, $N=20$, 95% confidence interval on r : (-0.14,0.78)),

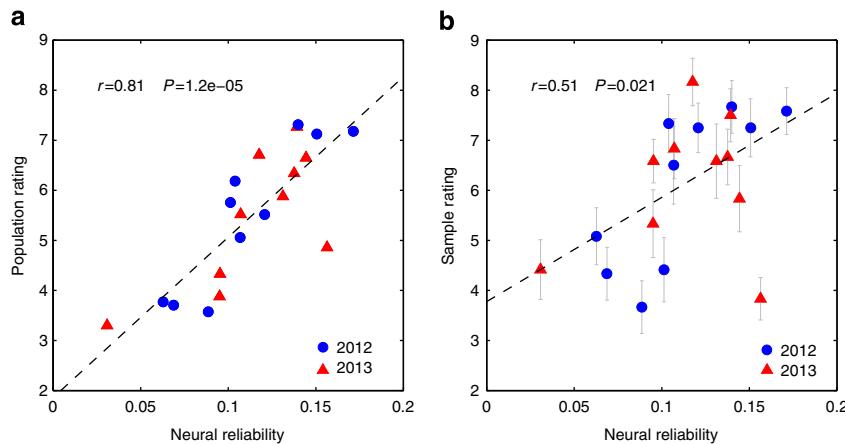


Figure 3 | Neural reliability in small sample is predictive of preference ratings in large audience. (a) Vertical axis: subjective ratings for 2×10 SuperBowl advertisements collected from a large online audience (Facebook-USA Today Ad Meter). Horizontal axis: neural reliability experienced across 12 subjects during each 30 s advertisement. Dashed line indicates the linear prediction of population ratings from neural reliability: 66% of variance in population ratings is explained. (b) Same as a, but with the vertical axis showing the mean ratings of the individuals in the sample (error bars indicate s.e.m., $N=12$). Only 26% of variance in the sample ratings is explained ($P=0.047$, $N=20$, Fisher r -to- z transformation).

which is significantly lower than the predictability of the population preferences ($P=0.047$, $N=20$, Fisher r -to- z transformation).

Could the reduced predictability of the sample ratings result from the variability due to the smaller sample size? To examine this, we generated $N=10^6$ random samples of 12 ratings for each of the 20 ads (assuming normal distributed ratings with the population rating as the mean and variance as observed in the actual sample). The resulting correlation of these simulated sample ratings with the neural reliability was significantly higher than what was observed for the actual sample ratings: a mean of $r=0.75$ with a 95% confidence interval of (0.66, 0.83), leading to a probability $P=4 \times 10^{-5}$ of drawing the actual value of $r=0.51$ from this distribution. We also explored the possibility of a systematic difference in the ratings of the sample and those of the population. However, ratings were largely consistent, differing significantly for only two of the 20 ads ($P>0.05$ false-discovery rate, $N=12$, Student t -test). A positive bias observed in the average rating (+0.65, $P=0.006$, $N=20$, Student t -test) should not affect correlation coefficients which are insensitive to a mean offset. Indeed, our sample ratings explain 59% of the variance in the population ratings ($r=0.77$, Supplementary Fig. 2), further lending credence to the notion that the reliability in neural responses is indeed more strongly linked to preferences of the population. Finally, it is worth noting that stimuli were judged with high preference heterogeneity: the same advertising was judged very differently by different subjects (the range of ratings for each ad was 6.25 ± 0.97 , that is, almost the full range of 1–10 was used by the 12 subjects to rate the ads).

To probe the spatial dimension of the observed neural reliability, we performed a follow-up fMRI experiment using a separate sample ($N=14$) of individuals, recording the BOLD signal evoked by all 20 of the SuperBowl ads. The subsequent BOLD activation time series were regressed onto the neural reliability scores (see horizontal axis of Fig. 3) in a block-design fashion. We sought to identify brain regions which exhibit systematically higher levels of activation for stimuli marked by high levels of neural reliability.

We found significant covariation of BOLD activity with EEG-derived neural reliability in both left and right lateral temporal cortices: these large clusters stretched from sensory association areas in occipital cortex, along the superior temporal gyrus, to the

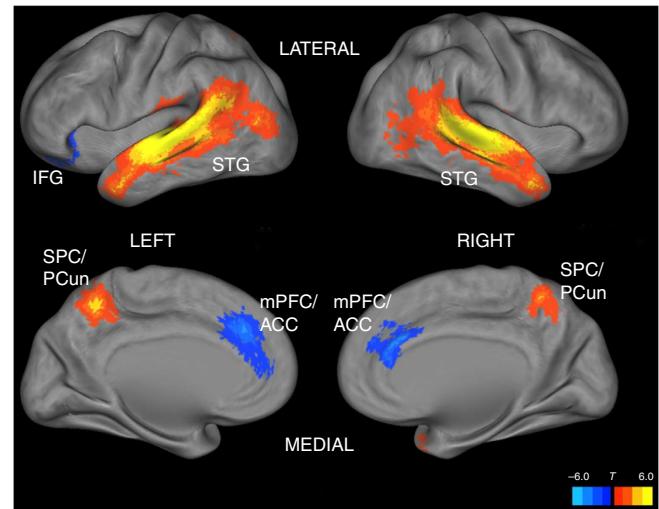


Figure 4 | Covariation of BOLD activity with EEG-derived neural reliability in different brain regions. Significant clusters of activation with a corrected family-wise error rate of 0.05 are mapped on inflated cortices using the CARET software⁵⁴. IFG, inferior frontal gyrus; STG, superior temporal gyrus; SPC/PCun, superior parietal cortex/precuneus; mPFC/ACC, medial prefrontal cortex/anterior cingulate cortex. Note that the BOLD and EEG data were collected from separate groups of subjects.

temporal poles (Fig. 4). Moreover, we observed significantly larger BOLD activation patterns for high-reliability advertisements in an area of parietal cortex including the superior parietal lobule and precuneus. Meanwhile, a significant negative covariation between neural reliability and BOLD activation was found in a region of medial prefrontal cortex (mPFC) that includes anterior cingulate cortex (ACC), as well as the left inferior frontal gyrus (IFG). To test if reliability of BOLD activity is also predictive of preference ratings, we computed the ISC of the spatiotemporal patterns of BOLD activity in the identified regions for each advertisement. The measured BOLD-ISC did not significantly correlate with the population nor the sample ratings ($r=0.34$ and $r=0.23$, respectively, $P>0.14$).

Discussion

Here, we showed that measures of behavioural responses aggregated over large audiences correlate significantly with the neural reliability evoked by the corresponding naturalistic stimuli in a small group of individuals. In particular, neural reliability is highly predictive of across-stimuli preferences, and predicts preferences of the large audience more accurately than those of the individuals from whom the neural activity was recorded.

Our finding differs subtly but importantly from those in which population responses are better predicted from a sample's neural activity than from its self-reports^{4,9}. Such findings may, in part, be explained by the fact that the behaviours of the population and sample are being evaluated with somewhat different measures (for example, expressing a preference for a stimulus versus actually consuming it). In the SuperBowl experiment described here, the behaviours performed by both the sample and population are identical, and their responses are well correlated. Note, however, that it is the population ratings that link most strongly to the reliability of neural responses, even though the sample is the source of the measured reliability. We have not found a precedent for the present observation that neural signals explain the population response better than the response of the sample.

One may conclude from the results that stimuli which evoke highly reliable neural responses among a small sample also do so in a larger audience. However, this interpretation does not account for the finding of significantly lower predictability of the sample ratings, which cannot be fully explained from the reduction in sample size. We conjecture that this finding is related to the preference heterogeneity of the advertising stimuli used: the high variability observed in the sample ratings may be attributed to differing subjective values^{2,22,23} or other variables such as social conformity^{24,25}. Such idiosyncratic processes may involve complex reasoning or emotional considerations that take relatively long to evaluate and presumably fail to yield immediate and reliable EEG signals. Through population aggregation, however, these idiosyncratic preferences tend to average out and one is left with what is shared by the large audience. Therefore, the surprising finding of this study is that reliability of relatively fast neural processing is a genuine predictor of the common preferences of a large population.

Preference heterogeneity has been studied extensively in the context of economic risk-taking, often focusing on the neural underpinnings of individual differences in decision-making²⁶. In the marketing literature, preference heterogeneity has been reported to affect perception of advertising^{27,28}. However, we are not aware of literature analysing the neural basis of preference heterogeneity with natural stimuli or, in particular, video advertisements.

Traditional neuroimaging work on the evaluation of preference or 'value' uses fMRI and points to elevated activity in specific subcortical regions²⁹. In particular, activity in the ventral striatum and medial prefrontal cortex (mPFC) correlates with individual subjective value²² and the purchasing behaviours of a larger population⁴. Such neural activity encodes information that is predictive of decisions following stimulus presentation^{7,30} even when measured in the absence of a choice³¹, thus pointing to a certain level of automatic stimulus evaluation. The present findings highlight the importance of reliable short-latency responses, suggesting similarly automatic stimulus processing. However, the present study points to neural processing of more superficial cortical areas, which are the main contributors to the EEG³². Note also that it is the reliability of temporal dynamics, and not necessarily the strength of response, that is carrying the predictive information here. It is also worthwhile to point out that previous efforts at analysing reliability of electrophysiological

signals required subdural electrodes and focused on slow modulations (in the order of 10 s) of oscillatory activity, in particular, the gamma band³³, which is known to correlate with the BOLD response³⁴. In contrast, here we used fast evoked responses measured on the scalp, which generally do not coincide with BOLD or gamma activity³⁵.

We observed that reliability of neural experience is related to subsequent preferences. However, note that in the Twitter and Nielsen studies, our behavioural measures index general response independent of valence; strictly speaking, the Tweet rates and Nielsen ratings are not reflective of 'liking' the stimulus, but rather being compelled to respond to or continue viewing it, respectively. Although it may be argued that tweeting about or tuning into a programme are behaviours consistent with 'liking' it, they are certainly not sufficient conditions for doing so. The present analysis has implicitly grouped both positive and negative valences into the dependent measure being predicted: for example, Twitter commentary to the episode expressed both positive as well as negative sentiment. It is thus possible that reliability correlates more generally with, for example, interest, rather than preference itself. On an anecdotal level, we do point out that the SuperBowl ad receiving the lowest population rating (unambiguously denoting a dislike) among all 20 ads also elicited the lowest neural reliability (see Supplementary Table 2).

It is interesting to contrast the present results with the literature on the neural basis of individual differences^{36–38}. There, the focus is on capturing neural features which vary across individuals and thus explain differences in individual behaviour. Here, we focus on the commonality in neural responses, effectively ignoring individual differences, to obtain a predictor of group behaviour. The component analysis technique used here to compute neural reliability explicitly looks for shared neural components, and the resulting quantity links closely to population measures which reflect shared behaviours (that is, trends) that emerge after aggregation of large samples.

The broad fMRI activations observed in sensory and association cortex suggest that modulations of high-level visual and auditory processing underlie the measured neural reliability. For example, the activations in bilateral temporal cortex may reflect processing of complex auditory speech information (both linguistic and prosodic) during advertisements³⁹. The activated region also included areas of occipitotemporal cortex recruited by the processing of dynamic visual stimuli^{40,41}. Increased BOLD activation was also found in the superior parietal lobules and precuneus, which mediates attention to auditory and visual stimuli^{42–44}. In addition, this region has also been associated with self-referential processing, imagery and memory⁴⁵, processes that may be elicited during the viewing of well-crafted advertising. Meanwhile, activations of the ACC/mPFC have been implicated in the evaluation of conflict and emotions⁴⁶, which may have occurred more frequently during the less likable advertisements. We caution, however, that all of these observed BOLD activations were found to co-vary with the EEG reliability of a separate group of subjects; as such, we refrain from inferring that the encephalographic signal components driving our preference-linked measure of neural reliability originate from these fMRI-identified regions. Although the topographies of the EEG components have been found to be fairly reproducible across various stimuli (see Fig. 5), the specific co-varying BOLD activations may be stimulus-dependent. Disparate neuro-modulatory processes may manifest in similar patterns of cortical generators which drive the observable EEG³².

It is possible that personal preferences yield changes in the individual's level of attention or engagement. Such 'top-down' modulation may then affect the strength⁴⁷ and thus the reliability of neural responses associated with stimulus-locked neural

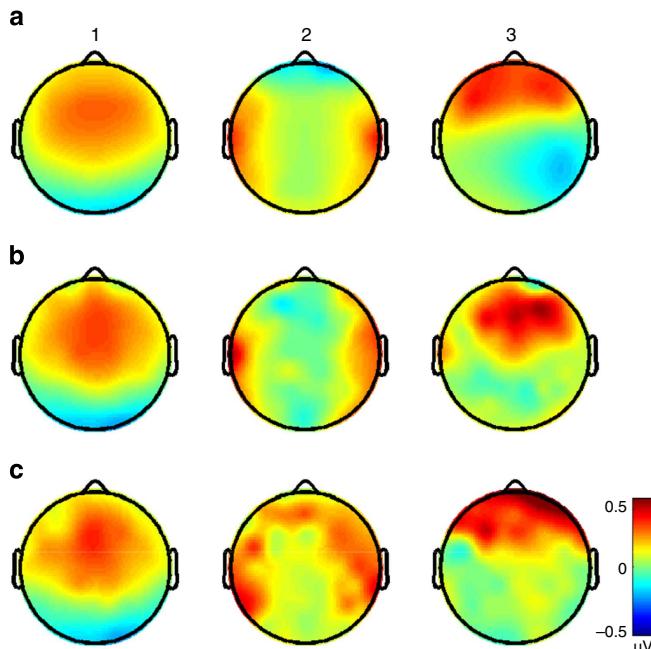


Figure 5 | Scalp projections of the three most-reliable dimensions of neural activity. (a) As measured during viewing of 'The Walking Dead' pilot, (b) as measured during viewing of 10 advertisements from the 2012 SuperBowl (c) same as b but from the 2013 SuperBowl.

processing. However, if individual preferences were to guide modulation of sensory processing, then we would have expected neural reliability to predict the sample preferences equally well, if not better, than the population preferences. Alternatively, it may be that individuals prefer stimuli precisely because the narratives drive the brain strongly and reliably. Such 'bottom-up' influence would evidently be well reflected in the preferences of large audiences; however, in the small sample, this sensory processing may be masked by the idiosyncratic preferences or biases of particular individuals.

The findings of the SuperBowl advertisement study suggest that stimuli which we judge favourably may be those to which our brains respond in a stereotypical manner that is shared by our peers. Viewed in another manner, if one is able to evoke reliable neural activity from one's audience, then that audience is, as a whole, more likely to find one's message favourable. However, the present data do not permit causal inference about the specific variables mediating the reliable patterns of activity. One possibility is that narrative elements of the stimuli directly bring about neural reliability. Indeed, disrupting the narrative structure for stimuli is known to reduce ISC for BOLD⁴⁸ and evoked responses¹⁶. But it is also possible that other aspects of the stimulus (for example, overall production quality) correlate with population preference⁴⁹, with this hidden variable explaining the link between advertisement ratings neural reliability. In this case, if one were to pinpoint the stimulus features that drive neural reliability, it would be possible to make the prediction of population behaviour directly from a content analysis of the stimulus (that is, without measuring neural responses).

Regardless of the source of the reliability-preference link, the finding that naturally occurring audience behaviours may be forecast from scalp measurements bears potentially tremendous relevance for fields outside the basic sciences such as education, marketing and media, which stand to gain from the predictive power of neural reliability.

Methods

Subjects and stimuli. For the encephalography recordings, 16 healthy subjects (nine females and seven males, ages 19–32, mean of 26 years) viewed the pilot episode of 'The Walking Dead' along with intermittent commercials as aired in the original broadcast. An additional 12 subjects (gender balanced, ages 20–29, mean of 25 years) viewed and subsequently rated (on a scale of 1–10) 10 advertisements initially aired during the 2012 SuperBowl (one subject was common to both experiments). To validate the results, the same subjects then viewed 10 ads from the 2013 SuperBowl. These 20 video clips were randomly selected and spanned the range of viewer ratings from the Facebook-USA Today Ad Meter (see Supplementary Table 2). For the fMRI recordings, a separate 14 subjects (six females, ages 18–22, mean of 20 years) viewed the same set of 2012 and 2013 SuperBowl advertisements. Subjects provided written informed consent in accordance with the procedures approved by the Institutional Review Boards of the City College of New York (EEG study) and the Georgia Institute of Technology (fMRI study).

EEG data collection. Study participants viewed the stimuli in a darkened, electrically and acoustically shielded room. Sound was played back with PC loudspeakers adjusted by each subject to a comfortable listening volume. Subjects were instructed to pay attention to the stimuli and to minimize overt movement. Before viewing, subjects were fitted with a 64-electrode cap placed on the scalp according to the international 10/10 standard for EEG, which was recorded with a BioSemi ActiveTwo system (BioSemi, Amsterdam, The Netherlands) at a sampling frequency of 512 Hz and 24 bits per sample. To subsequently correct eye-movement artifacts, we also recorded the electrooculogram (EOG) with four auxiliary electrodes (one adjacent to and one below each eye).

EEG preprocessing. All data processing was performed automatically (that is, with no manual intervention) offline in the MATLAB software (MathWorks, Natick, MA, USA). After extracting the EEG/EOG segments corresponding to the duration of each stimulus, the signals were high-pass filtered (1 Hz cutoff), notch filtered at 60 Hz, and down sampled to 256 Hz. Eye-movement related artifacts were corrected by linearly regressing out the four EOG channels from all EEG channels. The regression was performed on non-overlapping 5-s blocks for The Walking Dead data set, and on the entire data record for each SuperBowl advertisement (that is, a 30-s 'window'). After the correction of eye-movement artifacts, channels whose average power exceeded the mean channel power by four standard deviations were excluded from analysis, with this process repeated four times in an iterative scheme. Similarly, within each kept channel, samples whose squared-amplitude exceeded the mean-squared-amplitude of that channel by more than four standard deviations were rejected. Again, this procedure was iterated four times for each channel. In addition, we rejected every sample within 100 ms of the identified artifact samples. As our viewing paradigm did not constrain the subjects' eye movements during the relatively long stimulus durations, the data contained a larger proportion of artifacts than that seen in conventional, short-trial-based experiments. The proportion of data rejected for each scene of the Walking Dead episode is shown in Supplementary Fig. 3: there is no significant correlation between the time series and the log tweets per scene ($r = 0.04, P > 0.05$), nor between the time series and the prediction of log tweets per scene from neural reliability ($r = 0.002, P > 0.05$). Meanwhile, the proportion of data rejected for each SuperBowl ad is listed in Supplementary Table 2: there is no significant correlation between the proportion of data removed for each ad and the prediction of rating from the ISC ($r = 0.24, P > 0.05$), nor between the proportion of data removed for each ad and the population rating ($r = 0.21, P > 0.05$), nor between the proportion of data removed and the sample rating ($r = 0.09, P > 0.05$). In summary, the median (across subjects) percentage of samples removed was 15.98% for the Walking Dead data set, 16.28% for the 2012 SuperBowl data set and 19.95% for the 2013 SuperBowl data set. Rejected samples were marked as missing data ('NaN'), and the analysis proceeded by computing means and covariances with the *nanmean()* and *nancov()* MATLAB functions. As detailed in the next section, the method employed to compute reliability is rooted in covariance matrices whose sensitivity to outliers is well known; thus, we opted for a stringent outlier rejection to ensure robust covariance estimation.

Neural reliability computation. To compute the neural reliability elicited by a given stimulus, we employed the component analysis approach of Dmochowski *et al.*¹⁶, whose mathematical details are described below. The technique is similar to canonical correlation analysis⁵⁰ and its generalizations to multiple subjects⁵¹, differing in that it uses the same projection for all data sets. It is conceptually similar to the 'common canonical covariates' method⁵², which is based on a maximum-likelihood formulation, as opposed to the generalized eigenvalue problem developed in ref. 16.

For a given stimulus viewed by N subjects, we have a set of N data matrices $\{\mathbf{X}_1, \dots, \mathbf{X}_N\}$ where \mathbf{X}_n conveys the spatiotemporal neural response of subject n . We seek to project the data of all subjects onto a common space such that the resulting projections exhibit maximal ISCs across the subject pool. To that end, let $P_1 = \{p_{11}, p_{12}\} = \{(1,2), (1,3), \dots, (N-1, N)\}$ denote the set of all $P = N \times (N-1)/2$ unique

subject pairs. We then form the aggregated auto- and cross-covariance matrices as:

$$\begin{aligned}\mathbf{R}_{11} &= \frac{1}{PT} \sum_{i=1}^P \mathbf{X}_{p_{1i}} \mathbf{X}_{p_{1i}}^T \\ \mathbf{R}_{22} &= \frac{1}{PT} \sum_{i=1}^P \mathbf{X}_{p_{2i}} \mathbf{X}_{p_{2i}}^T \\ \mathbf{R}_{12} &= \frac{1}{PT} \sum_{i=1}^P \mathbf{X}_{p_{1i}} \mathbf{X}_{p_{2i}}^T,\end{aligned}\quad (1)$$

where T is the number of time samples (columns) in \mathbf{X}_n and T denotes matrix transposition.

We seek to find a projection vector \mathbf{w} which maximizes the ISC between subject-aggregated data records:

$$\frac{\mathbf{w}^T \mathbf{R}_{12} \mathbf{w}}{(\mathbf{w}^T \mathbf{R}_{11} \mathbf{w})^{1/2} (\mathbf{w}^T \mathbf{R}_{22} \mathbf{w})^{1/2}}\quad (2)$$

It is shown in ref. 16 that assuming $\mathbf{w}^T \mathbf{R}_{11} \mathbf{w} = \mathbf{w}^T \mathbf{R}_{22} \mathbf{w}$, the solution to equation (2) is a generalized eigenvalue problem:

$$\lambda(\mathbf{R}_{11} + \mathbf{R}_{22})\mathbf{w} = \mathbf{R}_{12}\mathbf{w},\quad (3)$$

where λ is the generalized eigenvalue corresponding the maximal ISC, encompassing all subject pairs, elicited by the stimulus. Note that the assumption $\mathbf{w}^T \mathbf{R}_{11} \mathbf{w} = \mathbf{w}^T \mathbf{R}_{22} \mathbf{w}$ does not limit generality, as one can simply define $p'_i = \{(1, 2), \dots, (N-1, N), (N, N-1), \dots, (2, 1)\}$ and then substitute p'_i in equation (1) to ensure that $\mathbf{R}_{11} = \mathbf{R}_{22}$; this was done in our analysis. Moreover, when computing the generalized eigenvalues of equation (3), we regularize the pooled auto-covariance by keeping only the first $K = 10$ dimensions. This value of K roughly corresponds to the ‘knee’ of the pooled auto-covariance eigenvalue spectrum in the spectral representation of $\mathbf{R}_{11} + \mathbf{R}_{22}$.

There are multiple non-orthogonal solutions to equation (3), whose associated generalized eigenvalues are ranked in decreasing order of aggregated ISC: $\lambda_1 > \lambda_2 > \dots > \lambda_D$, where D is the number of electrodes. We take the first $C = 3$ such solutions and linearly sum their corresponding eigenvalues to yield the estimate of

the population measure:

$$\text{predicted population response} = \beta_0 + \sum_{i=1}^C \beta_i \lambda_i,\quad (4)$$

where β_i is the regression coefficient relating the aggregated ISC in dimension i to the dependent population measure, as determined by linear least-squares. Note that due to the small sample size of the SuperBowl data set, the ISCs were uniformly summed across components to yield the estimate of neural reliability which was then directly correlated with the population measure:

$$\text{neural reliability} = \sum_{i=1}^C \lambda_i.\quad (5)$$

For The Walking Dead data set, we learned the optimal projections on data encompassing all scenes, and then applied these projections back onto the data of each scene to yield the time-resolved reliability in each component. In other words, the covariances in equation (1) were formed using data from all scenes, yielding the optimal \mathbf{w} , which was then applied to equation (2) but with the covariances there formed using only data for the desired segment of the stimulus. For the SuperBowl data set, we learned the optimal projections by concatenating the neural responses of all ads into a single data matrix per subject. Once again, this combined data was used to construct the covariance matrices and learn the optimal projection vectors. We then projected these optimized filters onto the data from each advertisement to compute the reliability exhibited by the participants’ during that ad.

Spatiotemporal characteristics of EEG components. Here we detail the spatial and temporal properties of the components formed from the optimal spatial filters \mathbf{w} . Let us construct a weight matrix \mathbf{W} whose columns represent the first C generalized eigenvectors in equation (3). The projections of the resulting components onto the scalp data are given by Parra *et al.*⁵³:

$$\mathbf{A} = \mathbf{RW}(\mathbf{W}^T \mathbf{RW})^{-1},\quad (6)$$

where $\mathbf{R} = \mathbf{R}_{11} + \mathbf{R}_{22}$ is the pooled auto-covariance. The columns of \mathbf{A} are termed ‘forward models’ and inform us of the approximate location of the underlying neuronal sources (up to the inherent limits imposed by volume conduction in EEG).

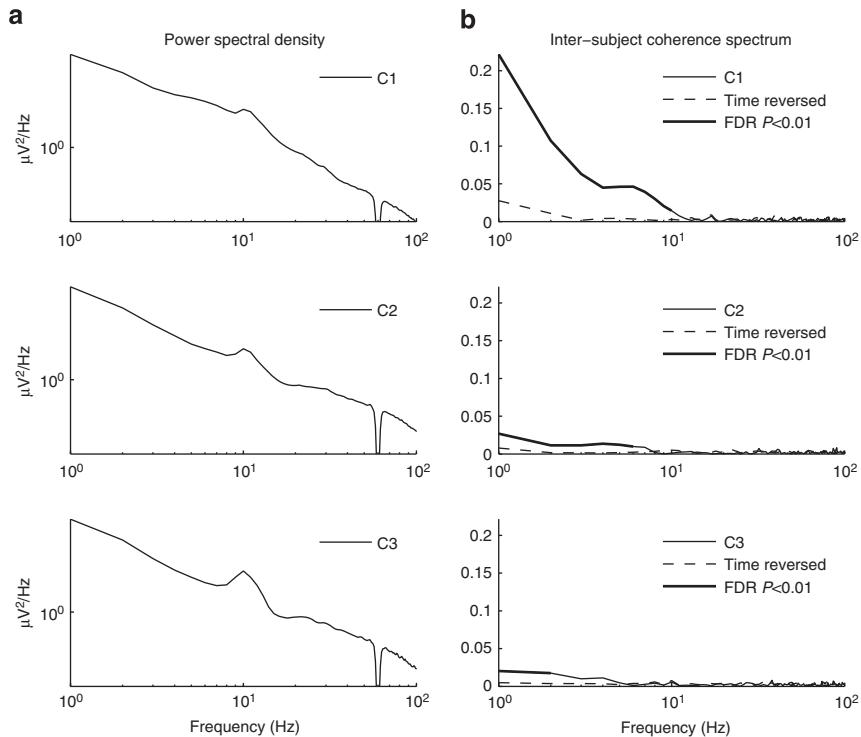


Figure 6 | Temporal properties of the signal components used to measure reliability. (a) Power-spectral density of each dimension of reliability averaged across all subjects from the 2012 SuperBowl data. A characteristic peak is evident for the alpha band (around 10 Hz). (b) Coherence spectrum computed for signals from pairs of subjects (66 unique pairs) and averaged across 10 video clips and all pairs (solid curve). Coherence is a frequency-resolved positive measure of correlation (on a scale of 0–1). Chance level of coherence was estimated by using the same signal pairs with one of the two reversed in time for the entire 30 s of each video clip (dashed curve). Significant difference was calculated using a bootstrap shuffle by pooling original and time-reversed coherence values and randomly drawing (10^5 times) among the pairs of signals. Significant coherence controlling for false-discovery rate at 0.01 indicated in bold.

Figure 5 depicts these forward models for the stimuli used in the study. The scalp projections stemming from The Walking Dead study bear a close resemblance to those found in ref. 16: a symmetric first component with a dipolar distribution consisting of frontocentral and occipital poles, a second component exhibiting bilateral poles at the temporal electrodes, and an asymmetric third component marked by frontal and right-parietal poles. Meanwhile, the forward models of the reliability-maximizing components from the SuperBowl study reveal a highly congruent first component topography, while deviating somewhat in the second and third components. For example, the frontal pole of the third component from the 2012 ads is slightly more posterior. Such disparities in scalp topographies may reflect a re-distribution of canonical sources among the three components.

Meanwhile, Fig. 6 summarizes the temporal properties of the components used to construct the measure of neural reliability. The 1/f temporal power spectrum of these components is typical for encephalography (Fig. 6a). A temporal coherence analysis of the signals used to measure neural reliability indicates that reliability is driven by immediate evoked responses of 2 s or less and can be as fast as 100 ms (Fig. 6b). Coherence across subjects—a frequency-resolved measure of correlation—is strongest at low frequencies, but statistically significant values can be found at frequencies as high as 10 Hz, consistent with previous findings using intra-cortical recordings³³.

fMRI data collection. For the fMRI recordings, we used the same two sets of ads from the 2012 and 2013 SuperBowls. All MRI data were acquired on a Siemens Magnetom Trio 3T scanner. A high-resolution T1 structural scan (3D MPRAGE, TI = 850 ms, flip angle = 9°, 1 mm isotropic resolution) was acquired before each subject viewed the ads. Before functional scanning, subjects were instructed to pay attention to the stimuli. Images were acquired using a whole-brain echo-planar imaging sequence (transverse orientation, TR = 2,000 ms, TE = 30 ms, flip angle = 90°, field of view = 204 mm) of 37 interleaved slices with 3 mm isotropic resolution and a 17% gap. Data were preprocessed to correct for slice timing to the first slice with a Fourier interpolation, using AFNI's 3dTshift tool. Head movements were then corrected using AFNI's 3dvolreg routine. Next, the functional data were smoothed with a 6 mm full-width half-maximum Gaussian kernel to reduce noise. Finally, data were transformed to the MNI standard space using FSL's FLIRT software using a 12-parameter trilinear affine transformation. The EEG reliability measure (see equation (5)) for each advertisement was used as an amplitude-modulated block-design regressor in a general linear model of the fMRI data including six motion parameters as covariates, using AFNI's 3dDeconvolve tool. Whole-brain group level analysis was performed using AFNI's 3dttest routine, with mixed effects inference on a one-sample *t*-test using individual beta values. AFNI's 3dClustSim tool was used with an estimated smoothing of 9.16 mm (obtained with AFNI's 3dFWHMx routine) to perform 10,000 Monte Carlo simulations to find a cluster size threshold (40 voxels) with a corrected family-wise error rate of 0.05; *P*(uncorrected) = 0.002. To compute the BOLD-ISC for each advertisement, we concatenated the BOLD time series of all significant voxels shown in Fig. 4, and then computed the correlation coefficient between all subject pairs. The resulting aggregated ISCs were then correlated with the population ratings.

Twitter data collection. Through the Crimson Hexagon service, we obtained a listing of all episode-related tweets which originated during the initial broadcast of The Walking Dead pilot, that is, all tweets from 10/31/2010 9:00–10:00 PM EST containing a relevant hashtag, referencing a show-specific Twitter account, or simply referencing the show's name. The listing was filtered to retain only those tweets which directly referenced episode content (that is, 1,947 of 19,000 total tweets). Each retained tweet was then manually linked to the corresponding scene(s) by inspecting the message content as well as the tweet timestamp. This procedure was performed by three research assistants who were blind to the hypothesis of the study. Approximately 61% of tweets could be unambiguously linked to one scene only; when ambiguous, the tweet was linked to multiple candidate scenes (that is, in general, the mapping from tweet to scene was one-to-many). The distribution of number of scenes referenced per tweet is shown in Supplementary Fig. 4. The tweet count of each scene was computed by summing the number of tweets referencing that scene. To subsequently analyse the relationship between neural reliability and Twitter reaction, we divided the tweet counts (incremented by one to handle forthcoming log operation) by scene duration to compensate for varying scene length, yielding a tweet-frequency. Finally, we logarithmically transformed the tweet rate to arrive at the dependent measure onto which neural reliability was regressed.

Nielsen data collection. Courtesy of AMC, we obtained the Nielsen ratings for each minute of the initial airing of pilot episode of The Walking Dead. We summed the ratings across age categories (18–49 and 25–54) and method of viewing (live versus digital video recorder). The regression results when predicting the viewer-ship within each age category are reported in Supplementary Table 1.

Ad Meter data collection. We obtained publicly available population-averaged scores for all 2012 and 2013 SuperBowl advertisements via the Facebook-USA Today Ad Meter service (Supplementary Table 2). An online panel of over 7,000 participants rated each advertisement on a scale of 1–5 for the 2012 commercials and on a scale of 1–10 for the 2013 commercials. We analysed each set separately and then combined

the results. For the combined analysis, the ratings of the 2012 commercials were linearly transformed under the assumption that the quality was comparable to that of 2013: the 2012 ratings were scaled and offset such that the entire set of ratings from 2012 matched the ratings from 2013 in mean and standard deviation.

References

- McClure, S. M. *et al.* Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* **44**, 379–387 (2004).
- O'Doherty, J. P., Buchanan, T. W., Seymour, B. & Dolan, R. J. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* **49**, 157 (2006).
- Plassmann, H., O'Doherty, J., Shiv, B. & Rangel, A. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl Acad. Sci. USA* **105**, 1050–1054 (2008).
- Berns, G. S. & Moore, S. E. A neural predictor of cultural popularity. *J. Consum. Psychol.* **22**, 154–160 (2012).
- Falk, E. B., O'Donnell, M. B. & Lieberman, M. D. Getting the word out: neural correlates of enthusiastic message propagation. *Front. Hum. Neurosci.* **6**, 313 (2012).
- Levy, D. J. & Glimcher, P. W. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* **31**, 14693–14707 (2011).
- Tusche, A., Bode, S. & Haynes, J. D. Neural responses to unattended products predict later consumer choices. *J. Neurosci.* **30**, 8024–8031 (2010).
- Falk, E. B., Berkman, E. T., Mann, T., Harrison, B. & Lieberman, M. D. Predicting persuasion-induced behavior change from the brain. *J. Neurosci.* **30**, 8421–8424 (2010).
- Falk, E. B., Berkman, E. T. & Lieberman, M. D. From neural responses to population behavior neural focus group predicts population-level media effects. *Psychol. Sci.* **23**, 439–445 (2012).
- Ariely, D. & Berns, G. S. Neuromarketing: the hope and hype of neuroimaging in business. *Nat. Rev. Neurosci.* **11**, 284–292 (2010).
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. & Malach, R. Intersubject synchronization of cortical activity during natural vision. *Science* **303**, 1634–1640 (2004).
- Hasson, U., Furman, O., Clark, D., Dudai, Y. & Davachi, L. Enhanced intersubject correlations during movie viewing correlate with successful episodic encoding. *Neuron* **57**, 452–462 (2008).
- Stephens, G. J., Silbert, L. J. & Hasson, U. Speaker-listener neural coupling underlies successful communication. *Proc. Natl Acad. Sci. USA* **107**, 14425–14430 (2010).
- Nummenmaa, L. *et al.* Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl Acad. Sci. USA* **109**, 9599–9604 (2012).
- Hanson, S. J., Gagliardi, A. D. & Hanson, C. Solving the brain synchrony eigenvalue problem: conservation of temporal dynamics (fmri) over subjects doing the same task. *J. Comput. Neurosci.* **27**, 103–114 (2009).
- Dmochowski, J. P., Sajda, P., Dias, J. & Parra, L. C. Correlated components of ongoing EEG point to emotionally laden attention—a possible marker of engagement? *Front. Hum. Neurosci.* **6**, 112 (2012).
- Lankinen, K. *et al.* Intersubject consistency of cortical MEG signals during movie viewing. *NeuroImage* **92**, 217–224 (2014).
- Sasahara, K., Hirata, Y., Toyoda, M., Kitsuregawa, M. & Aihara, K. Quantifying collective attention from tweet stream. *PLoS ONE* **8**, e61823 (2013).
- Fisher, R. A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* **10**, 507–521 (1915).
- Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap*, Vol. 57 (CRC press, 1994).
- Hasson, U. *et al.* Neurocinematics: the neuroscience of film. *Projections* **2**, 1–26 (2008).
- Kable, J. W. & Glimcher, P. W. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* **10**, 1625–1633 (2007).
- Peters, J. & Büchel, C. Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *J. Neurosci.* **29**, 15727–15734 (2009).
- Izuma, K. The neural basis of social influence and attitude change. *Curr. Opin. Neurobiol.* **23**, 456–462 (2013).
- Rilling, J. K. & Sanfey, A. G. The neuroscience of social decision-making. *Annu. Rev. Psychol.* **62**, 23–48 (2011).
- Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T. & Platt, M. L. Neural signatures of economic preferences for risk and ambiguity. *Neuron* **49**, 765–776 (2006).
- Horsky, D., Misra, S. & Nelson, P. Observed and unobserved preference heterogeneity in brand-choice models. *Market. Sci.* **25**, 322–335 (2006).
- Feick, L. & Higie, R. A. The effects of preference heterogeneity and source characteristics on ad processing and judgements about endorsers. *J. Advertising* **21**, 9–24 (1992).

29. Lebreton, M., Jorge, S., Michel, V., Thirion, B. & Pessiglione, M. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* **64**, 431–439 (2009).
30. Ma, Y., Wang, C. & Han, S. Neural responses to perceived pain in others predict real-life monetary donations in different socioeconomic contexts. *Neuroimage* **57**, 1273–1280 (2011).
31. Levy, I., Lazzaro, S. C., Rutledge, R. B. & Glimcher, P. W. Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J. Neurosci.* **31**, 118–125 (2011).
32. Nunez, P. L. *Electric Fields of the Brain: The Neurophysics of EEG* (Oxford University Press, 2006).
33. Honey, C. J. *et al.* Slow cortical dynamics and the accumulation of information over long timescales. *Neuron* **76**, 423–434 (2012).
34. Lachaux, J. P. *et al.* Relationship between task-related gamma oscillations and bold signal: new insights from combined fmri and intracranial eeg. *Hum. Brain Mapp.* **28**, 1368–1375 (2007).
35. Foucher, J. R., Otzenberger, H. & Gounot, D. The bold response and the gamma oscillations respond differently than evoked potentials: an interleaved eeg-fmri study. *BMC Neurosci.* **4**, 22 (2003).
36. Mohr, P. N. & Nagel, I. E. Variability in brain activity as an individual difference measure in neuroscience? *J. Neurosci.* **30**, 7755–7757 (2010).
37. Hamann, S. & Canil, T. Individual differences in emotion processing. *Curr. Opin. Neurobiol.* **14**, 233–238 (2004).
38. Berkman, E. T. & Falk, E. B. Falk Beyond brain mapping: using neural measures to predict real-world outcomes. *Curr. Dir. Psychol. Sci.* **22**, 45–50 (2013).
39. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
40. Ebisch, S. J. *et al.* Differential involvement of somatosensory and interoceptive cortices during the observation of affective touch. *J. Cogn. Neurosci.* **23**, 1808–1822 (2011).
41. Petit, L. & Haxby, J. V. Functional anatomy of pursuit eye movements in humans as revealed by fMRI. *J. Neurophysiol.* **82**, 463–471 (1999).
42. Nardo, D., Santangelo, V. & Macaluso, E. Stimulus-driven orienting of visuo-spatial attention in complex dynamic environments. *Neuron* **69**, 1015–1028 (2011).
43. Shomstein, S. & Yantis, S. Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *J. Neurosci.* **26**, 435–439 (2006).
44. Wise, S. P., Boussaoud, D., Johnson, P. B. & Caminiti, R. Premotor and parietal cortex: Corticocortical connectivity and combinatorial computations 1. *Annu. Rev. Neurosci.* **20**, 25–42 (1997).
45. Cavanna, A. E. & Trimble, M. R. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* **129**, 564–583 (2006).
46. Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* **15**, 85–93 (2011).
47. Hopfinger, J. B. & West, V. M. Interactions between endogenous and exogenous attention on cortical visual processing. *Neuroimage* **31**, 774–789 (2006).
48. Hasson, U., Malach, R. & Heeger, D. J. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* **14**, 40 (2010).
49. Prag, J. & Casavant, J. An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry. *J. Cult. Econ.* **18**, 217–235 (1994).
50. Hotelling, H. Relations between two sets of variates. *Biometrika* **28**, 321–377 (1936).
51. Kettenring, J. R. Canonical analysis of several sets of variables. *Biometrika* **58**, 433–451 (1971).
52. Neuenschwander, B. E. & Flury, B. D. Common canonical covariates. *Biometrika* **82**, 553–560 (1995).
53. Parra, L. C., Spence, C. D., Gerson, A. D. & Sajda, P. Recipes for the linear analysis of EEG. *Neuroimage* **28**, 326–341 (2005).
54. Van Essen, D. C. A population-average, landmark-and surface-based (pals) atlas of human cerebral cortex. *Neuroimage* **28**, 635–662 (2005).

Acknowledgements

We thank the following individuals: Joanna Raczkiewicz and Clint Beharry for proposing, together with B.P.A., the use of the Walking Dead episode, and, together with Kelly Creighton, Andrew Bowe and Claris Chang, developing the coding schemes for collecting and labelling Twitter data; Jason Sherwin for establishing the collaboration between the CCNY and the HI; Victoria Romero for establishing the collaboration between CCNY and GA Tech; and Tony Norcia, Paul Sajda, Ofer Tchernichovski, Davide Reato and Simon Kelly for reading and providing feedback on earlier versions of the manuscript; Uri Hasson and his students for fruitful discussions; David Fein for help with the neural recordings. Initial brainstorming meetings that led to the Walking Dead analysis included J.P.D., B.P.A., J.S.J., Jason Sherwin, Joanna Raczkiewicz, Clint Beharry and Paul Sajda. We also acknowledge the support from the Defense Advanced Research Projects Agency (government grant numbers D13AP00003 and D12PC00397).

Author contributions

J.P.D. and L.C.P. designed EEG experiments, analysed the data and wrote the paper. M.A.B., E.H.S. and L.C.P. designed the fMRI experiment and wrote the corresponding portion of the paper. J.P.D. collected and processed the EEG data. M.A.B. collected and analysed the fMRI data. B.P.A. compiled the Twitter data, proposed the use of and compiled the Nielsen data. J.S.J. facilitated access to the Twitter and Nielsen data sets.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Dmochowski, J. P. *et al.* Audience preferences are predicted by temporal reliability of neural processing. *Nat. Commun.* **5**:4567 doi: 10.1038/ncomms5567 (2014).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Anatomically distinct dopamine release during anticipation and experience of peak emotion to music

Valorie N Salimpoor^{1–3}, Mitchel Benovoy^{3,4}, Kevin Larcher¹, Alain Dagher¹ & Robert J Zatorre^{1–3}

Music, an abstract stimulus, can arouse feelings of euphoria and craving, similar to tangible rewards that involve the striatal dopaminergic system. Using the neurochemical specificity of [¹¹C]raclopride positron emission tomography scanning, combined with psychophysiological measures of autonomic nervous system activity, we found endogenous dopamine release in the striatum at peak emotional arousal during music listening. To examine the time course of dopamine release, we used functional magnetic resonance imaging with the same stimuli and listeners, and found a functional dissociation: the caudate was more involved during the anticipation and the nucleus accumbens was more involved during the experience of peak emotional responses to music. These results indicate that intense pleasure in response to music can lead to dopamine release in the striatal system. Notably, the anticipation of an abstract reward can result in dopamine release in an anatomical pathway distinct from that associated with the peak pleasure itself. Our results help to explain why music is of such high value across all human societies.

Humans experience intense pleasure to certain stimuli, such as food, psychoactive drugs and money; these rewards are largely mediated by dopaminergic activity in the mesolimbic system, which has been implicated in reinforcement and motivation (see ref. 1 for a review). These rewarding stimuli are either biological reinforcers that are necessary for survival, synthetic chemicals that directly promote dopaminergic neurotransmission, or tangible items that are secondary rewards. However, humans have the ability to obtain pleasure from more abstract stimuli, such as music and art, which are not directly essential for survival and cannot be considered to be secondary or conditioned reinforcers. These stimuli have persisted through cultures and generations and are pre-eminent in most people's lives. Notably, the experience of pleasure to these abstract stimuli is highly specific to cultural and personal preferences, which can vary tremendously across individuals.

Most people agree that music is an especially potent pleasurable stimulus² that is frequently used to affect emotional states. It has been empirically demonstrated that music can effectively elicit highly pleasurable emotional responses^{3,4} and previous neuroimaging studies have implicated emotion and reward circuits of the brain during pleasurable music listening^{5–8}, particularly the ventral striatum^{5–7}, suggesting the possible involvement of dopaminergic mechanisms⁹. However, the role of dopamine has never been directly tested. We used ligand-based positron emission tomography (PET) scanning to estimate dopamine release specifically in the striatum on the basis of the competition between endogenous dopamine and [¹¹C]raclopride for binding to dopamine D₂ receptors¹⁰. Pleasure is a subjective phenomenon that is difficult to assess objectively. However, physiological changes occur during moments of extreme pleasure, which can be used to index pleasurable states in response to music. We used the 'chills' or 'musical frisson'¹¹ response, a well-established marker of

peak emotional responses to music^{5,12–14}. Chills involve a clear and discrete pattern of autonomic nervous system (ANS) arousal¹⁵, which allows for objective verification through psychophysiological measurements. Thus, the chills response can be used to objectively index pleasure, a subjective phenomenon that would otherwise be difficult to operationalize, and allows us to pinpoint the precise time of maximal pleasure.

Previous studies have typically used experimenter-selected musical stimuli^{6–8}. However, musical preferences are highly individualized; thus, to ensure maximal emotional responses, participants were asked to select their own highly pleasurable music. After extensive screening (Online Methods), we recruited a group of people who consistently experienced objectively verifiable chills during their peak emotional responses so that we could quantify both the occurrence and the timing of the most intense pleasurable responses. We also collected psychophysiological measurements (heart rate, respiration rate, electrodermal skin conductance, blood volume pulse amplitude and peripheral temperature) during the PET scans to verify ANS differences between conditions. To account for psychoacoustical differences across self-selected stimuli, we matched musical excerpts using a previously established procedure⁵, such that participants listened to one another's choices, which served as either pleasurable or neutral stimuli. We predicted that if the rewarding aspects of music listening are mediated by dopamine, substantial [¹¹C]raclopride binding potential differences would be found between neutral and pleasurable conditions in mesolimbic regions.

The second aim of our study was to explore the temporal dynamics of any dopaminergic activity, as distinct anatomical circuits are thought to underlie specific phases of reward responses^{16,17}. That is, if there is dopamine release, we wanted to examine whether it is associated with the experience of the reward or with its anticipation¹⁸. Music provides

¹Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada. ²International Laboratory for Brain, Music and Sound Research, Montreal, Quebec, Canada. ³Centre for Interdisciplinary Research in Music Media and Technology, Montreal, Quebec, Canada. ⁴Centre for Intelligent Machines, McGill University, Montreal, Quebec, Canada. Correspondence should be addressed to V.N.S. (valorie.salimpoor@mail.mcgill.ca) or R.J.Z. (robert.zatorre@mcgill.ca).

Received 7 October 2010; accepted 25 November 2010; published online 9 January 2011; doi:10.1038/nn.2726

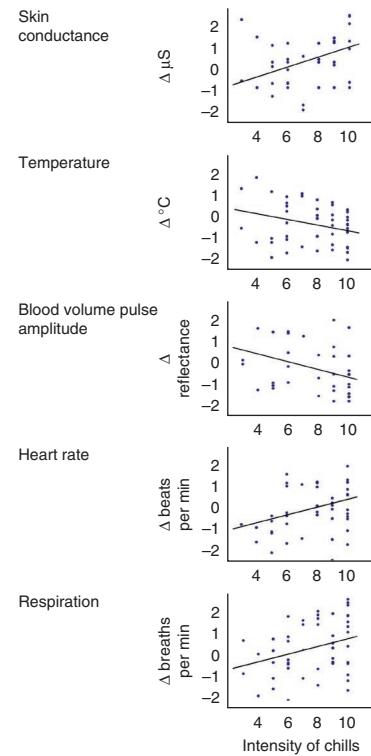
Figure 1 Positive correlation between emotional arousal and intensity of chills during PET scanning. The mean intensity of chills reported by each participant during the PET scanning session was significantly correlated with psychophysiological measurements that were also acquired during the scan. These are indicative of increased sympathetic nervous system activity, suggesting that the intensity of chills is a good marker of peak emotional arousal (**Supplementary Table 1**). The y axis represents standardized z scores for each biosignal. See main text for *P*-values.

an innovative means of assessing this distinction because the temporal unveiling of tonal arrangements elicits anticipatory responses that are based on cognitive expectations and prediction cues^{11,19,20}. These can be examined to isolate the functional components that precede peak pleasurable responses. As PET does not afford the temporal resolution required to examine this distinction, we combined the temporal specificity of functional magnetic resonance imaging (fMRI) with the neurochemical specificity of PET. We acquired fMRI scans with the same participants and stimuli to examine the temporal profile of blood oxygenation level-(BOLD) response specifically in those regions that also showed dopamine release with PET. Striatal dopamine release and BOLD responses are known to be correlated, although the relationship is complex^{9,21}. We predicted that regions revealing dopamine activity in the PET data would show the largest increases in hemodynamic response during peak emotional experiences. We separately analyzed the BOLD data from epochs of peak pleasure and the time immediately preceding these responses (that is, anticipation), based on participants' real-time behavioral responses of when chills were experienced. Spatial conjunction analyses were used to confine the analysis to those striatal voxels showing both dopamine release from PET and increased BOLD during fMRI, which ensured that we were measuring the hemodynamic signal only from regions known to release dopamine in response to the same stimuli. This multimodal procedure revealed a temporally mediated distinction in dopamine release to anticipatory and consummatory responses in the dorsal and ventral striatum, respectively.

RESULTS

PET data: dopamine release and emotional arousal

PET scanning took place over two sessions. Participants listened to either pleasurable music or neutral music during the entire session while both subjective and objective indicators of emotional arousal were collected. Subjective responses from rating scales included self-reports of number of chills, intensity of chills and degree of pleasure



experienced from each excerpt. The mean number of chills for each pleasurable music excerpt was 3.7 (s.d. = 2.8). A paired-samples *t* test confirmed that greater pleasure was experienced during the pleasurable music condition over the neutral music condition ($t(49) = 25.0$, $P < 0.001$). Notably, there was a significant positive correlation between the reported intensity of chills and the reported degree of pleasure ($r = 0.71$, $P < 0.001$), suggesting that the chills response is a good representation of pleasure experienced amongst this group.

Objective measures of psychophysiological signals indicative of emotional arousal collected during the two PET scanning sessions showed significantly higher ANS activity during the pleasurable music condition in all of the variables that we measured: namely, increases in heart rate ($P < 0.05$), respiration ($P < 0.001$) and electrodermal response ($P < 0.05$), and decreases in temperature ($P < 0.01$) and blood volume pulse amplitude ($P < 0.001$; for values, see (**Supplementary Table 1**)). Subjective reports of the intensity of the chills response collected via rating scales during PET scanning were significantly correlated with the degree of ANS arousal on all measures: increases in heart rate ($P < 0.05$), respiration ($P < 0.05$) and electrodermal response ($P < 0.01$), and decreases

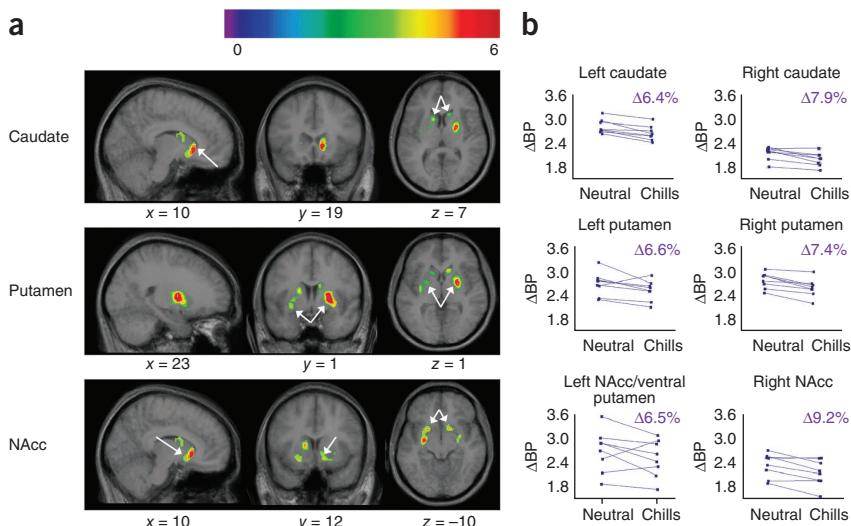
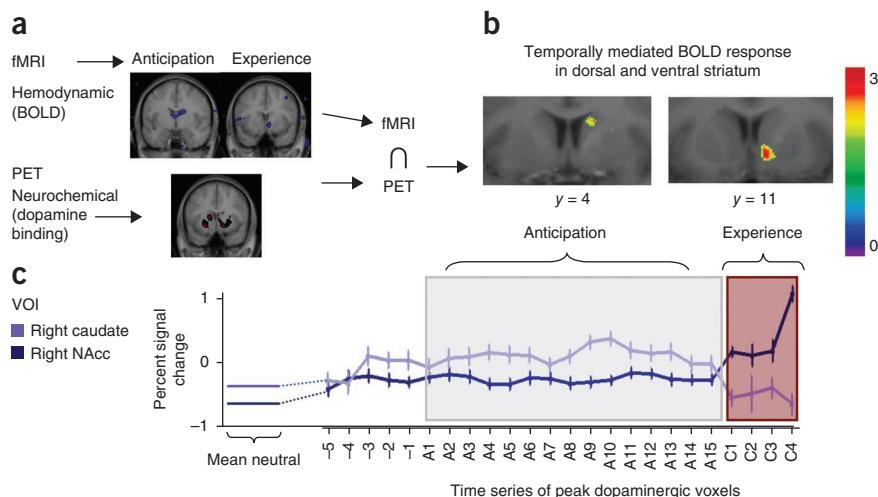


Figure 2 Evidence for dopamine release during pleasurable music listening. **(a)** Statistical parametric maps (*t* statistic on sagittal, coronal and axial slices) reveal significant ($P < 0.001$) [^{11}C]raclopride binding potential (BP) decreases bilaterally in the caudate, putamen and NAcc (white arrows) during pleasurable compared with neutral music listening (**Supplementary Table 2**), indicating increased dopamine release during pleasurable music. **(b)** Changes in binding potential (BP) values plotted separately for each individual; note that the change was consistent for the majority of people at each site.

Figure 3 Combined fMRI and PET results reveal temporal distinctions in regions showing dopamine release. (a) $[^{11}\text{C}]$ raclopride PET scan results were spatially conjoined with the fMRI results by creating a mask of significant dopamine release overlayed on BOLD response t maps during each condition. (b) Hemodynamic responses and dopamine activity were maximal in the caudate during anticipatory phases, but shifted more ventrally to NAcc during peak emotional responses. (c) Percent signal change in BOLD response relative to the mean was calculated from the peak voxel of the caudate and NAcc clusters based on the $[^{11}\text{C}]$ raclopride PET data. Voxels showing maximum dopamine release in the caudate and NAcc (**Supplementary Table 2**) were identified and percent BOLD signal change was calculated during the fMRI epochs associated with peak emotional responses; values were interpolated for each second preceding this response for each individual, up to 15 s, which was defined as the anticipatory period based on previous findings¹⁵ (see Online Methods for additional details). We found increased activity during anticipation (A1-A15) and decreased activity during peak emotional response (C1-C4) for the caudate, but a continuous increase in activity in NAcc with a maximum during peak emotional responses. The mean signal for neutral epochs for the NAcc and caudate clusters are also plotted for reference, as are the 5 s preceding the anticipation epochs.



in temperature ($P < 0.05$) and blood volume pulse amplitude ($P < 0.05$; (Fig. 1 and **Supplementary Table 1**). This finding further verified that the chills response is a good objective representation of peak emotional arousal in this group.

Analysis of PET data (**Supplementary Methods**) revealed increased endogenous dopamine transmission, as indexed by decreases in $[^{11}\text{C}]$ raclopride binding potential, bilaterally in both the dorsal and ventral striatum ($P < 0.001$; Fig. 2a) when contrasting the pleasurable music with the neutral music condition. The percentage of dopamine binding potential change was highest in the right caudate and the right nucleus accumbens (NAcc; Fig. 2b and **Supplementary Table 2**). These results indicate that the experience of pleasure while listening to music is associated with dopamine release in striatal reward systems.

fMRI data: temporal specificity of reward responses

To gain information about the dynamics of dopamine release over time, we acquired fMRI scans during presentation of pleasurable and neutral music excerpts. Listeners indicated by button press when they experienced chills (mean = 3.1 chills per excerpt, s.d. = 0.9); these responses

were then used *post hoc* to identify anticipation and peak experience time periods (Fig. 3a). Anticipation epochs were defined as 15 s before the peak experiences. BOLD responses for each of these epochs were compared with periods in which participants reported feeling neutral during the same musical excerpts. The result of this contrast for each of the events was then spatially conjoined with a mask of regions that had released dopamine according to the $[^{11}\text{C}]$ raclopride PET scan. We found that hemodynamic activity in the regions showing dopamine release was not constant throughout the excerpt, but was restricted to moments before and during chills and, critically, was anatomically distinct. During peak pleasure experience epochs, as compared with neutral epochs, there was increased BOLD response in the right NAcc ($x, y, z = 8, 10, -8$; $t = 2.8$; Fig. 3b). In contrast, increased BOLD response was also found during the anticipation epochs, but was largely confined to the right caudate ($x, y, z = 14, -6, 20$; $t = 3.2$; Fig. 3b).

The temporal dynamics of the reward response and its relationship to the caudate and NAcc clusters can be more specifically analyzed by examining the percent BOLD signal change occurring over time in relation to peak pleasure. To avoid the ‘circularity’ problem²², we

Figure 4 Brain and behavior relationships involving temporal components of pleasure during music listening. Left, coronal slices showing binding potential differences in dorsal (top) and ventral (bottom) striatum that also show hemodynamic activity during anticipation versus experience of chills, respectively. Right, behavioral ratings of the number and intensity of chills and pleasure reported during the PET scans plotted against $[^{11}\text{C}]$ raclopride binding potential changes in the two clusters. The number of chills reported was positively correlated with percent binding potential change in the caudate (* $P < 0.05$), which was linked to BOLD response immediately preceding chills (that is, anticipatory periods), consistent with the idea that a greater number of chills would result in greater anticipation and result in more activity in the areas associated with anticipation. The mean intensity of chills and reported pleasure were positively correlated with the NAcc (** $P < 0.01$), which was linked to BOLD response during chills, confirming that this region is involved in the experience of the highly pleasurable component of music listening.

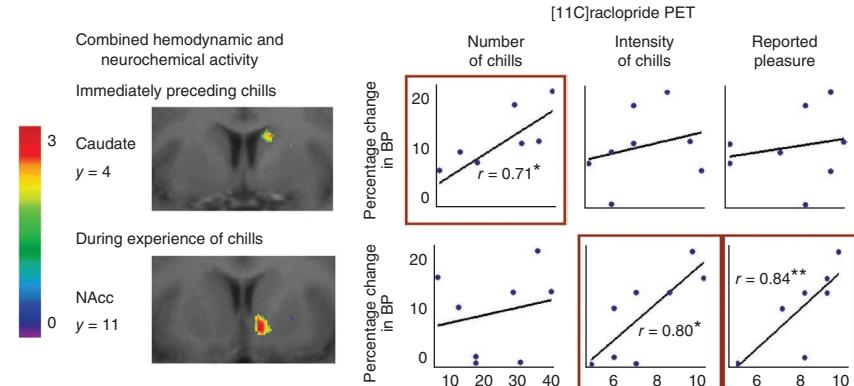


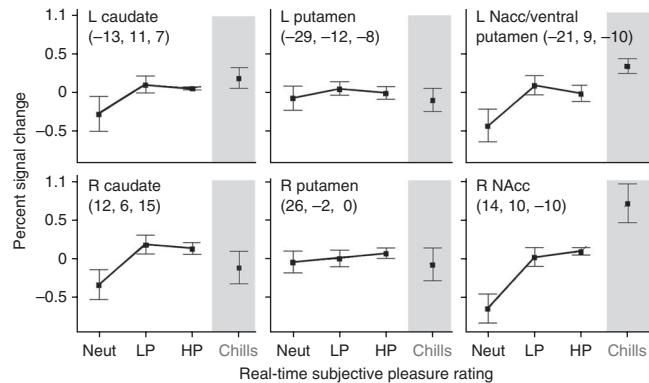
Figure 5 Brain and behavior relationships involving parametric increases in pleasure during music listening. Relationship between real-time ratings of pleasure during music listening and percent BOLD signal change relative to the mean in regions showing dopamine release as identified via PET. The chills epochs (shaded) were excluded from the analysis (values shown here only for reference) to examine activity related to increases in pleasure irrespective of chills. A regression analysis revealed that the NAcc, and to a lesser extent the left and right caudate, significantly predicted increases in pleasure ratings during each of the conditions ($P < 0.05$ and $P < 0.001$, respectively; **Supplementary Table 4**). This analysis indicates that activity in these regions increased with pleasure even when no chills were experienced. LP, low pleasure; HP, high pleasure.

derived our voxels of interest (VOIs) from the PET data, which are independent of the fMRI data. This procedure also allowed us to better integrate the hemodynamic and neurochemical results. We found that activity in both the caudate and NAcc was increased during anticipation as compared with the mean signal during the neutral epochs for the same pieces of music, with larger increases occurring in the caudate (Fig. 3c). During the peak emotional response, however, activity in the caudate decreased, whereas activity in the NAcc continued to increase. These findings support our fMRI contrast results and provide temporal information as to how hemodynamic activity in the regions showing dopamine release may contribute to reward processing in real time.

Brain-behavior relationships

Once we had identified, via fMRI, the caudate and NAcc as contributing to the anticipation and experience, respectively, of peak pleasure moments during music listening, we used our PET scan data to further explore the brain and behavior relationships in these clusters. Mean $[^{11}\text{C}]$ raclopride binding potential values from the NAcc and caudate clusters were plotted against behavioral data obtained during PET scanning, which required participants to indicate the total number of chills, mean intensity of chills and mean subjective pleasure experienced during each piece of music. We found that the number of chills was significantly correlated ($P < 0.05$) with binding potential differences in the right caudate, but not the NAcc, whereas the intensity of chills and overall degree of pleasure experienced were most significantly correlated ($P < 0.01$) with binding potential change in the right NAcc, but not the caudate (Fig. 4 and **Supplementary Table 3**). This finding further supports a functional dissociation in the contribution of these anatomical regions to pleasure associated with music listening.

An additional question is whether increases in pleasure alone, in the absence of chills, result in increased hemodynamic responses in the same areas as during the experience of chills, although perhaps not to the same extent. We examined this question by determining whether there was a linear relationship between increases in pleasure and hemodynamic activity in the right NAcc, irrespective of chills, and how this compared with other striatal regions showing dopamine release. This analysis was done by excluding all of the epochs during which individuals experienced chills and examining BOLD signal changes that related to increasing pleasure in the right NAcc. Using the voxel that showed the maximum dopamine release in the NAcc during the $[^{11}\text{C}]$ raclopride scan, we calculated the percent BOLD signal change as subjective pleasure ratings increased from neutral to low pleasure to high pleasure (excluding chills) for each individual. Note that this analysis, unlike the one presented above, does not take into account the temporal component, as all epochs rated as having the same pleasure were averaged, regardless of when they occurred with respect to chills. A regression analysis



revealed a significant linear trend in which the percent signal change in the right NAcc accounted for 67% of the variability in subjective pleasure ratings ($t(19) = 6.18, P < 0.001$). This finding suggests that increases in subjective pleasure correspond to increases in neural activity in the NAcc, in the same regions as those involved in the chills responses and those that showed dopamine release in the PET study, even though this analysis excluded all chills epochs.

Next, to ensure that increases in pleasure, irrespective of chills, are not better predicted by activity in regions of the striatum other than the right NAcc, we performed a similar analysis in all anatomical clusters that had shown dopamine release in the PET study. We first selected peak voxels from each cluster showing dopamine release from the PET data and then extracted the percent BOLD signal change as listeners reported increases in pleasure from the fMRI data; as before, all chills epochs were excluded. A stepwise multiple regression analysis was performed to examine which cluster's hemodynamic responses were best able to predict pleasure states. We found that hemodynamic increase in the NAcc cluster was the most significant predictor ($P < 0.01$) of increasing subjective pleasure (**Supplementary Table 4**). However, at a lower statistical threshold of $P < 0.05$, bilateral caudate clusters and the left NAcc/ventral putamen cluster could also predict pleasure states, but to a lower degree (31% and 43% for left and right caudate, respectively, and 37% for the NAcc/ventral putamen). Recruitment of the caudate is not surprising considering that anticipatory periods result in a culmination of pleasurable emotional experiences and the caudate was recruited during these pleasant anticipatory moments. Indeed, the mean subjective pleasure rating provided by listeners during the anticipatory epochs was 2.51 (s.d. = 0.55), which was significantly higher than that of the entire excerpt (mean = 2.11, s.d. = 0.019; $t(246) = 8.5, P < 0.001$).

Finally, when the percent BOLD signal change during the chills epochs was included in the multiple regression analysis (Fig. 5), it was apparent that the experience of chills represents the highest point of hemodynamic activity in the NAcc. These findings converge to suggest that the dorsal and ventral subdivisions of the striatum are most involved during anticipation and experience of the peak emotional responses during music listening, respectively.

DISCUSSION

Our results provide, to the best of our knowledge, the first direct evidence that the intense pleasure experienced when listening to music is associated with dopamine activity in the mesolimbic reward system, including both dorsal and ventral striatum. This phylogenetically ancient circuitry has evolved to reinforce basic biological behaviors with high adaptive value. However, the rewarding qualities of music listening are not obviously directly adaptive. That is, musical stimuli,

similar to other aesthetic stimuli, are perceived as being rewarding by the listener, rather than exerting a direct biological or chemical influence. Furthermore, the perception that results in a rewarding response is relatively specific to the listener, as there is large variability in musical preferences amongst individuals. Thus, through complex cognitive mechanisms, humans are able to obtain pleasure from music², a highly abstract reward consisting of just a sequence of tones unfolding over time, which is comparable to the pleasure experienced from more basic biological stimuli.

One explanation for this phenomenon is that it is related to enhancement of emotions^{3,15,20}. The emotions induced by music are evoked, among other things, by temporal phenomena, such as expectations, delay, tension, resolution, prediction, surprise and anticipation^{11,19}. Indeed, we found a temporal dissociation between distinct regions of the striatum while listening to pleasurable music. The combined psychophysiological, neurochemical and hemodynamic procedure that we used revealed that peaks of ANS activity that reflect the experience of the most intense emotional moments are associated with dopamine release in the NAcc. This region has been implicated in the euphoric component of psychostimulants such as cocaine²³ and is highly interconnected with limbic regions that mediate emotional responses, such as the amygdala, hippocampus, cingulate and ventromedial prefrontal cortex²⁴. In contrast, immediately before the climax of emotional responses there was evidence for relatively greater dopamine activity in the caudate. This subregion of the striatum is interconnected with sensory, motor and associative regions of the brain^{24,25} and has been typically implicated in learning of stimulus-response associations^{24,26} and in mediating the reinforcing qualities of rewarding stimuli such as food²⁷. Our findings indicate that a sense of emotional expectation, prediction and anticipation in response to abstract pleasure can also result in dopamine release, but primarily in the dorsal striatum. Previous studies have found that amphetamine-induced dopamine release in the NAcc spreads to more dorsal regions after repeated exposure to the drug²⁸, which suggests that this area may be involved in improved predictability and anticipation of a reward. Similarly, previous studies involving rewards such as food and smoking that contain a number of contextual predicting cues (for example, odor and taste) also found dorsal striatum dopamine release^{27,29}. Conversely, in studies in which there were no contextual cues or experience with the drugs involved, dopamine release was largely observed in the ventral striatum^{30,31}. Finally, evidence from animal research also suggests that, as rewards become better predicted, the responses that initiated in the ventral regions move more dorsally in the striatum³². These results are consistent with a model in which repeated exposure to rewards associated with a specific context gradually shift the response from ventral to dorsal and further suggest that contextual cues that allow prediction of a reward, in our case the sequences of tones leading up to the peak pleasure moments, may also act as reward predictors mediated via the dorsal striatum.

Another noteworthy finding is the correspondence between behavioral and imaging results, which strengthens the evidence for the distinct roles of dorsal and ventral striatum. We found a positive correlation between subject-reported intensity of chills and dopamine release in the NAcc during [¹¹C]raclopride PET scanning (Fig. 4), which confirms the fMRI results that peak pleasure responses are associated with this region. Furthermore, the number of chills reported by listeners during the PET scan was correlated with dopamine release in the caudate (Fig. 4), which is consistent with the fMRI results showing increased activity in this region during anticipation of peak emotional responses; as greater number of chills suggests increased incidence of anticipation, greater dopamine release would be expected in this area.

It is important to note that chills are not necessarily pleasurable *per se*, as they can be unpleasant in other contexts (for example, as a result of intense fear). Instead, chills are physiological markers of intense ANS arousal^{5,15,33,34}, which in turn is believed to underlie peak pleasure during music listening^{5,15}; we used chills here only to allow objective quantification of a highly subjective response that would be otherwise difficult to measure and because they afford precision as to the time at which the peak pleasure occurred. As such, chills are byproducts, and not a cause of the emotional responses. Thus, it is important to clarify that, although chills index peak emotional responses in this group of people, the specific experience of chills is not necessary to result in neural activity in the striatum, a finding that is consistent with less-specific analyses performed in previous studies^{6–8}. This conclusion is confirmed by our findings that, even when the chills epochs were excluded from the analysis, there was still a significant linear relationship between increases in self-reported pleasure and increases in hemodynamic activity in the regions that showed dopamine release (Fig. 5). Furthermore, when chills were reported, maximal signal was seen in the NAcc voxels that showed a linear increase as participants progressed from neutral to low pleasure to high pleasure, further confirming that chills represented the peak of pleasure in this group. This finding is also consistent with the finding that the degree of binding potential decrease in the NAcc for each participant was positively correlated with the degree of pleasure reported from listening to the musical excerpts, irrespective of the number of chills that were experienced (Fig. 4).

It should be noted that there was some activity in the ventral striatum during the anticipation phase at lower statistical thresholds, consistent with other studies using different stimuli²⁴. However, we found that, during the anticipatory phase, there was also increased BOLD response in the caudate (more so than the NAcc), which then shifted more ventromedially as participants reported experiencing peak reward (Fig. 3). This is an important finding because the stimulus that we are using is a dynamic reward with a temporal component, allowing examination of the reward in real time as it progresses from anticipation to peak pleasure states, which is generally not possible because of limitations with movement inside the PET scanner. Some studies administered the pleasurable stimulus (for example, food) immediately before the scan and measured subsequent dopamine release²⁷, in which case anticipation and consumption cannot be distinguished. Other studies measured the anticipation phase online, with the promise of the delivery of the tangible reward after the scan, in which case the consumption phase is missed^{35,36}. Music is a unique reward that allows assessment of all reward phases online, from the point that a single note is heard to the point at which maximum pleasure is reached.

The anatomical dissociation between the anticipatory and consummatory phases during intensely pleasurable music listening suggests that distinct mechanisms are involved. This distinction may map onto the ‘wanting’ and ‘liking’ phases of a reward in an error prediction model³⁷. The anticipatory phase, set off by temporal cues signaling that a potentially pleasurable auditory sequence is coming, can trigger expectations of euphoric emotional states and create a sense of wanting and reward prediction. This reward is entirely abstract and may involve such factors as suspended expectations and a sense of resolution. Indeed, composers and performers frequently take advantage of such phenomena, and manipulate emotional arousal by violating expectations in certain ways or by delaying the predicted outcome (for example, by inserting unexpected notes or slowing tempo) before the resolution to heighten the motivation for completion. The peak emotional response evoked by hearing the desired sequence would

represent the consummatory or liking phase, representing fulfilled expectations and accurate reward prediction. We propose that each of these phases may involve dopamine release, but in different subcircuits of the striatum, which have different connectivity and functional roles.

The notion that dopamine can be released in anticipation of an abstract reward (a series of tones) has important implications for understanding how music has become pleasurable. However, the precise source of the anticipation requires further investigation. A sense of anticipation may arise through one's familiarity with the rules that underlie musical structure, such that listeners are anticipating the next note that may violate or confirm their expectations, in turn leading to emotional arousal, or alternatively it may arise through familiarity with a specific piece and knowing that a particularly pleasant section is coming up¹¹. These components are not mutually exclusive, as the second likely evolves from the first, and the overall anticipation is likely to be a combination of both. Nonetheless, the subtle differences that exist between them will need to be disentangled through future experiments that are specifically designed to parse out this distinction. Abstract rewards are largely cognitive in nature and our results pave the way for future work to examine nontangible rewards that humans consider rewarding for complex reasons.

Dopamine is pivotal for establishing and maintaining behavior. If music-induced emotional states can lead to dopamine release, as our findings indicate, it may begin to explain why musical experiences are so valued. These results further speak to why music can be effectively used in rituals, marketing or film to manipulate hedonic states. Our findings provide neurochemical evidence that intense emotional responses to music involve ancient reward circuitry and serve as a starting point for more detailed investigations of the biological substrates that underlie abstract forms of pleasure.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/natureneuroscience/>.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We thank the staff of the Montreal Neurological Institute PET and MR Units and the staff of the Centre for Interdisciplinary Research in Music Media and Technology for help with data acquisition, M. Ferreira and M. Bouffard for their assistance with data analysis, and G. Longo for assistance with stimulus preparation. This research was supported by funding from the Canadian Institutes of Health Research to R.J.Z., a Natural Science and Engineering Research Council stipend to V.N.S., a Jeanne Timmins Costello award to V.N.S. and Centre for Interdisciplinary Research in Music Media and Technology awards to V.N.S. and M.B.

AUTHOR CONTRIBUTIONS

V.N.S., R.J.Z. and A.D. designed the study. V.N.S. and M.B. performed all experiments. V.N.S., M.B. and K.L. analyzed the data. V.N.S. and R.J.Z. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/natureneuroscience/>.

Reprints and permissions information is available online at <http://www.nature.com/reprintsandpermissions/>.

- Egerton, A. *et al.* The dopaminergic basis of human behaviors: a review of molecular imaging studies. *Neurosci. Biobehav. Rev.* **33**, 1109–1132 (2009).
- Dube, L. & Lebel, J. The content and structure of laypeople's concept of pleasure. *Cogn. Emot.* **17**, 263–295 (2003).
- Sloboda, J. & Juslin, P.N. Psychological perspectives on music and emotion. in *Music and Emotion: Theory and Research* (ed. Sloboda, J.) 71–104 (Oxford University Press, Oxford, 2001).
- Krumhansl, C.L. An exploratory study of musical emotions and psychophysiology. *Can. J. Exp. Psychol.* **51**, 336–353 (1997).
- Blood, A.J. & Zatorre, R.J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. USA* **98**, 11818–11823 (2001).
- Menon, V. & Levitin, D.J. The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage* **28**, 175–184 (2005).
- Koelsch, S., Fritz, T., Crimon, D., Muller, K. & Friederici, A.D. Investigating emotion with music: an fMRI study. *Hum. Brain Mapp.* **27**, 239–250 (2006).
- Mitterschiffthaler, M.T., Fu, C.H.Y., Dalton, J., Andrew, C.M. & Williams, S. A functional MRI study of happy and sad affective states induced by classical music. *Hum. Brain Mapp.* **28**, 1150–1162 (2007).
- Knutson, B. & Gibbs, S.E. Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology (Berl.)* **191**, 813–822 (2007).
- Laruelle, M. Imaging synaptic neurotransmission with *in vivo* binding competition techniques: a critical review. *J. Cereb. Blood Flow Metab.* **20**, 423–451 (2000).
- Huron, D. & Hellmuth Margulis, E. Musical expectancy and thrills. in *Music and Emotion* (eds. Juslin, P.N. & Sloboda, J.) (Oxford University Press, New York, 2009).
- Grewé, O., Nagel, F., Kopiez, R. & Altenmüller, E. Emotions over time: synchronicity and development of subjective, physiological, and facial affective reactions to music. *Emotion* **7**, 774–788 (2007).
- Panksepp, J. The emotional source of "chills" induced by music. *Music Percept.* **13**, 171–207 (1995).
- Sloboda, J. Music structure and emotional response: some empirical findings. *Psychol. Music* **19**, 110–120 (1991).
- Salimpoor, V.N., Benovoy, M., Longo, G., Cooperstock, J.R. & Zatorre, R.J. The rewarding aspects of music listening are related to degree of emotional arousal. *PLoS ONE* **4**, e7487 (2009).
- O'Doherty, J.P., Deichmann, R., Critchley, H.D. & Dolan, R. Neural responses during anticipation of a primary taste reward. *Neuron* **33**, 815–826 (2002).
- Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
- Wise, R.A. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494 (2004).
- Huron, D. *Sweet Anticipation: Music and the Psychology of Expectation* (MIT Press, Cambridge, Massachusetts, 2006).
- Meyer, L.B. *Emotion and Meaning in Music*. (University of Chicago Press, Chicago, 1956).
- Schott, B.H. *et al.* Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *J. Neurosci.* **28**, 14311–14319 (2008).
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S. & Baker, C.I. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
- Volkow, N.D. *et al.* Relationship between subjective effects of cocaine and dopamine transporter occupancy. *Nature* **386**, 827–830 (1997).
- Haber, S. & Knutson, B. The reward circuit: linking primate anatomy and human imaging. *Neuropharmacology* **55**, 4–26 (2010).
- Haber, S.N., Kim, K.S., Mailly, P. & Calzavarini, R. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J. Neurosci.* **26**, 8368–8376 (2006).
- Valentin, V.V. & O'Doherty, J.P. Overlapping prediction errors in dorsal striatum during instrumental learning with juice and money reward in the human brain. *J. Neurophysiol.* **102**, 3384–3391 (2009).
- Small, D.M., Jones-Gotman, M. & Dagher, A. Feeding-induced dopamine release in dorsal striatum correlates with meal pleasantness ratings in healthy human volunteers. *Neuroimage* **19**, 1709–1715 (2003).
- Boileau, I. *et al.* Modeling sensitization to stimulants in humans: an [¹¹C]raclopride/positron emission tomography study in healthy men. *Arch. Gen. Psychiatry* **63**, 1386–1395 (2006).
- Barrett, S.P., Boileau, I., Okker, J., Pihi, R.O. & Dagher, A. The hedonic response to cigarette smoking is proportional to dopamine release in the human striatum as measured by positron emission tomography and [¹¹C]raclopride. *Synapse* **54**, 65–71 (2004).
- Leyton, M. *et al.* Amphetamine-induced increases in extracellular dopamine, drug wanting, and novelty seeking: a PET/[¹¹C]raclopride study in healthy men. *Neuropsychopharmacology* **27**, 1027–1035 (2002).
- Boileau, I. *et al.* Alcohol promotes dopamine release in the human nucleus accumbens. *Synapse* **49**, 226–231 (2003).
- Everitt, B.J. & Robbins, T. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* **8**, 1481–1489 (2005).
- Rickard, N.S. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychol. Music* **32**, 371–388 (2004).
- Grewé, O., Kopiez, R. & Altenmüller, E. Chills as an indicator of individual emotional peaks. *Ann. NY Acad. Sci.* **1169**, 351–354 (2009).
- Zald, D.H. *et al.* Dopamine transmission in the human striatum during monetary reward tasks. *J. Neurosci.* **24**, 4105–4112 (2004).
- Koepf, M.J. *et al.* Evidence for striatal dopamine release during a video game. *Nature* **393**, 266–268 (1998).
- Zald, D.H. & Zatorre, R.J. On music and reward. in *The Neurobiology of Sensation and Reward* (ed. Gottfried, J.A.) (CRC Press, 2011).

ONLINE METHODS

Participant screening and stimulus selection. 217 individuals responded to advertisements requesting people who experience chills to music; after five rounds of screening, the final group included eight participants. First, individuals provided ten pieces of instrumental music to which they experience intense pleasure and “chills” without restrictions to the genre of music, which included classical, folk, jazz, electronica, rock, punk, techno and tango (see http://www.zlab.mcgill.ca/supplements/supplements_intro.html for samples). Next, an email questionnaire was completed to determine whether their chills were experienced at times of extreme pleasure, consistently at the same point in the music without diminishing on multiple listening, in different environments, and the selected music was not specifically or generally associated with an episodic memory. 45 individuals continued to the third screening session, where a history of medical, psychiatric illness, or substance abuse was ruled out. 40 participants continued to the fourth screening session, where control stimuli were selected for each individual using a paradigm where one individual’s pleasurable music is used as another person’s neutral music^{5,15}. This way, group-averaged data analysis involves comparison of similar sets of stimuli. Although we were not able to match perfectly between the control and pleasurable pieces used for all participants, efforts were made to ensure that pieces were as evenly distributed as possible. Each individual rated other participants’ music on a scale of 1–10 (neutral to extremely pleasurable). From the pieces rated neutral, the ones that were most familiar to that subject were selected to minimize differences in familiarity between pleasurable music and neutral music conditions. Individuals whose music was found to be “neutral” by at least one other participant were asked to continue. Participants were asked not to listen to those pieces anymore during the course of the study to ensure maximal responses during testing. 28 individuals participated in the final screening session to verify the chills response at prespecified times through subjective and physiological responses. Participants listened to their chills-inducing music while providing subjective ratings of pleasure through button presses and indicating when they experienced a chill (see ref. 15 for additional details). The ten participants (five female, five male) who most reliably experienced chills during their peak pleasure responses to music accompanied by clear increases in ANS activity were selected for the study. The final group of participants was between the ages of 19 and 24 ($M = 20.8, \pm 1.9$ years) and had a wide range of musical experiences from no training to 15 years of experience.

Procedures. Ethical approval for the study was granted by the Montreal Neurological Institute (MNI) Research Ethics Board. All individuals gave written informed consent before participating in the study. Testing took place over three sessions. The first two sessions involved PET scanning (Supplementary Fig. 1) and psychophysiological recording (Supplementary Fig. 2) and the third session involved fMRI scanning (Supplementary Fig. 3).

Statistical analysis. Signal filtering was performed to remove noise and artifacts (see ref. 15 for additional details). Data were downsampled to 1-s epochs and compared across neutral music and pleasurable music conditions. To account for unequal variances across conditions, we used Welch’s *t* test. A second analysis was performed to examine the relationship between the intensity of chills experienced and psychophysiological responses. Outliers beyond four s.d. from the mean were removed for each excerpt and for each participant individually (2–5% of the data points). Subjective ratings for one individual were not recorded and BVP amplitude data for one participant demonstrated excessive artifacts, thus these data were not included in the analysis. Z score values of each biosignal were calculated for each excerpt and plotted against subjective ratings of chills intensity that subjects reported after hearing each excerpt (Fig. 1). Correlation coefficients were calculated for the intensity of chills and changes in each of the psychophysiological measures (Supplementary Table 1).

For [¹¹C]raclopride PET, we discarded two datasets because of participant discomfort during the first session. Data from the remaining eight participants were analyzed. PET emission frames were reconstructed and corrected for gamma ray attenuation and scatter. All PET images were corrected for head motion using a co-registration-based method, which performs interframe realignment and compensates for emission-transmission mismatches³⁸. The motion-corrected PET data were summed over the time dimension and aligned to the subject’s anatomical magnetic resonance image. Anatomical MRI were transformed into standardized stereotaxic space by means of automated feature

matching algorithm to the MNI template³⁹. All transformed images were visually inspected to ensure that there were no alignment errors.

Parametric images were generated in the native PET space by computing [¹¹C]raclopride binding potential (binding potential = B_{Avail} / K_D , where B_{Avail} is the density of available receptors and K_D is the dissociation constant) at each voxel of interest^{40,41}. Voxelwise [¹¹C]raclopride binding potential was calculated using a simplified reference region method^{40,41}, with the cerebellum chosen as reference region because it does not contain specific D₂ receptor-like binding sites and can be used for the determination of nonspecific binding and free radioligand in the brain⁴². The gray matter of the cerebellum assigned as reference region was initially segmented in Talairach space from a probabilistic atlas⁴³ and a neural net classifier⁴⁴. The [¹¹C]raclopride binding potential maps were then transformed into MNI space³⁹ using the previously determined transformation parameters. Statistical parametric *t* maps of binding potential change were produced by comparing the parametric binding potential maps of the two scan sessions (pleasurable music and neutral music), using a previously described method⁴⁵. This calculation uses the residuals of the least-squares fit of the compartmental model, which improves the sensitivity to small changes by providing better estimates of the s.d. at the voxel and by increasing the degrees of freedom. It is assumed that a reduction in [¹¹C]raclopride binding potential is indicative of an increase in extracellular dopamine concentration⁴⁶. Clusters of significant change were defined as all contiguous striatal voxels on the *t* map exceeding a magnitude threshold of 3.11. This threshold was considered to be significant ($P < 0.05$, corrected for multiple comparisons) for a search volume equal to the striatum and an effective spatial resolution of 8-mm full-width at half maximum (FWHM)⁴⁷. Mean binding potential values were extracted from each significant cluster for each individual and percent change in binding potential was calculated as $[(BP_{\text{neutral}} - BP_{\text{pleasurable}}) \times 100 / BP_{\text{neutral}}]$, and compared with subjectively reported post-listening ratings of the number of chills, intensity of chills and degree of pleasure experienced.

fMRI. One scan was terminated because of claustrophobia. fMRI data were corrected for motion using in-house software. To increase the signal-to-noise ratio, we spatially smoothed the images (or low-pass filtered) with an 8-mm FWHM isotropic Gaussian kernel. Image analyses were performed with fMRISTAT, which consists of a series of MATLAB scripts that utilize the general linear model for analyses⁴⁸. The general linear model ($Y = X\beta + \epsilon$) expresses the response variable (BOLD signal) Y in terms of a linear combination of explanatory variables (events) X , the parameter estimates (effects of interest) β and the error term ϵ . Temporal drift was modeled as cubic splines and removed by inclusion into the general linear model as a variable of non-interest. The linear model was solved for the parameter estimates β with least squares, yielding estimates of effects, standard errors and *t* statistics for each contrast and for each run.

Before group statistical maps for each contrast of interest were generated, in-house software was used to linearly transform anatomical and functional images from each subject into standard MNI stereotaxic coordinate space using the MNI 305 template³⁹. A mixed-effects linear model was subsequently used to combine data across subjects; the s.d. images were smoothed with a Gaussian filter so that the ratio of the random-effects variance divided by the fixed-effects variance results in approximately 100 degrees of freedom. Because the main purpose of the fMRI analyses was to measure BOLD activity in predefined striatal regions, we adopted an uncorrected statistical threshold of $P < 0.01$.

For the main analysis, three events were defined: the peak emotional response (PER) condition represented all epochs during which the participant was pressing the chills, the anticipation condition represented 15-s epochs immediately preceding the onset of the PER condition defined *post hoc*, and the neutral condition represented all epochs during which participants were pressing down the neutral button. Note that these neutral epochs are different from the neutral music condition, which were not used in this case, as the neutral music condition contrasted with the pleasurable music condition shows less activity in the striatum. As such, any epoch selected from the pleasurable music condition, even those not related to peak pleasure, could have shown increased striatal activity and overestimated the results of the study. The anticipation period was defined as the 15 s before the PER based on previous findings that this is the time frame during which psychophysiological responses begin to increase significantly relative to mean responses throughout the excerpt¹⁵. The times at which participants pressed the low pleasure and high pleasure buttons were also included in the model to ensure

that they did not contribute to baseline. A 0.1-s epoch was incorporated into the model each time a button was pressed to account for neural activity involved in button pressing. The BOLD data from times when participants were responding to questions were excluded from the analysis. The planned comparisons for the main analysis were then entered into the analysis: anticipation of PER = anticipation condition minus neutral condition and experience of PER = PER condition minus neutral condition.

Time series analysis. To further investigate the temporal dynamics of the reward response, we calculated the time series of hemodynamic activity in the caudate and NAcc clusters. To avoid the circularity problem²², we derived our VOIs from the PET data, which are independent of the fMRI data. We first identified the voxel showing the maximum dopamine release during the [¹¹C]raclopride PET scan, in the caudate and NAcc clusters. We then extracted the mean signal for each VOI during the entire fMRI run obtained from each volume and calculated the percent BOLD signal change relative to the mean of the run during the epochs in which PERs were reported. Participants often experienced multiple chills one after another. For the purposes of this analysis, the percent signal change during the first chill of the series was used, which ranged in duration from 1–4 s. The BOLD response for each of those seconds is plotted in **Figure 3c**. Mean signal change for each second preceding this response for each individual, up to 15 s, was also plotted to demonstrate hemodynamic time series during the anticipation period. As a result of cardiac gating, a different number of frames were acquired for each person during this 15-s period and acquisition time varied from 2.1 to 3 s depending on the individual's heart rate. As such, the VOI values obtained at each frame were interpolated to provide an estimate of signal during each second preceding the peak response. The mean number of frames sampled for calculating time series was 5.3 (s.d. = 1.3) during anticipation and 1.6 (s.d. = 1.2) during chills. The mean signal change during neutral button presses was also calculated for each VOI separately and plotted in **Figure 3d** for reference. Finally, the percent signal change for 5 s preceding the anticipatory response were also plotted for reference.

Conjunction analysis. Because [¹¹C]raclopride binds with D2 receptors mainly in the striatum⁴⁹, our fMRI data analysis was also limited to this region, masked

by areas that showed dopamine release. A spatial conjunction analysis was performed to examine the temporal aspects of hemodynamic activity in areas that had shown changes in [¹¹C]raclopride binding potential on PET. A mask of striatal areas that had revealed substantial changes in binding potential using the stated threshold ($t \geq 3.11$) was created to spatially mask both contrasts (outlined in the fMRI data analysis section): anticipation of PER and experience of PER. This procedure allowed us to measure BOLD changes only in voxels that had shown binding potential differences in the PET study.

38. Costes, N. *et al.* Motion correction of multi-frame PET data in neuroreceptor mapping: simulation based validation. *Neuroimage* **47**, 1496–1505 (2009).
39. Collins, D.L., Peters, T.M. & Evans, A.C. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J. Comput. Assist. Tomogr.* **18**, 192–205 (1994).
40. Gunn, R.N., Lammertsma, A.A., Hume, S.P. & Cunningham, V.J. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *Neuroimage* **6**, 279–287 (1997).
41. Lammertsma, A.A. & Hume, S.P. Simplified reference tissue model for PET receptor studies. *Neuroimage* **4**, 153–158 (1996).
42. Litton, J.E., Hall, H. & Pauli, S. Saturation analysis in PET-analysis of errors due to non-perfect reference regions. *J. Cereb. Blood Flow Metab.* **14**, 358–361 (1994).
43. Collins, D.L. & Evans, A.C. ANIMAL: Validation and application of nonlinear registration-based segmentation. *Intern. J. Pattern Recognit. Artif. Intell.* **11**, 1271–1294 (1997).
44. Zijdenbos, A., Forghani, R. & Evans, A.C. Automatic quantification of MS lesions in 3D MRI Brain data sets: validation of INSECT. in *Medical Image Computing and Computer-Assisted Intervention* (eds. Wells, W.M., Colchester, A. & Delp, S.) 439–448 (Springer-Verlag, Cambridge, Massachusetts, 1998).
45. Aston, J.A. *et al.* A statistical method for the analysis of positron emission tomography neuroreceptor ligand data. *Neuroimage* **12**, 245–256 (2000).
46. Endres, C.J. *et al.* Kinetic modeling of [¹¹C]raclopride: combined PET microdialysis studies. *J. Cereb. Blood Flow Metab.* **17**, 932–942 (1997).
47. Worsley, K.J. *et al.* A unified statistical approach for determining significant signals in images of cerebral activation. *Hum. Brain Mapp.* **4**, 58–73 (1996).
48. Worsley, K.J. *et al.* A general statistical analysis for fMRI data. *Neuroimage* **15**, 1–15 (2002).
49. Slifstein, M. *et al.* Striatal and extrastriatal dopamine release measured with PET and [(18)F]fallypride. *Synapse* **64**, 350–362 (2010).

Marketing actions can modulate neural representations of experienced pleasantness

Hilke Plassmann*, John O'Doherty*, Baba Shiv†, and Antonio Rangel**

*Division of the Humanities and Social Sciences, California Institute of Technology, MC 228-77, Pasadena, CA 91125; and †Stanford Graduate School of Business, Stanford University, 518 Memorial Way, Littlefield L383, Stanford, CA94305

Edited by Leslie G. Ungerleider, National Institutes of Health, Bethesda, MD, and approved December 3, 2007 (received for review July 24, 2007)

Despite the importance and pervasiveness of marketing, almost nothing is known about the neural mechanisms through which it affects decisions made by individuals. We propose that marketing actions, such as changes in the price of a product, can affect neural representations of experienced pleasantness. We tested this hypothesis by scanning human subjects using functional MRI while they tasted wines that, contrary to reality, they believed to be different and sold at different prices. Our results show that increasing the price of a wine increases subjective reports of flavor pleasantness as well as blood-oxygen-level-dependent activity in medial orbitofrontal cortex, an area that is widely thought to encode for experienced pleasantness during experiential tasks. The paper provides evidence for the ability of marketing actions to modulate neural correlates of experienced pleasantness and for the mechanisms through which the effect operates.

orbitofrontal cortex | modulation by marketing actions | neuroeconomics | taste

A basic assumption in economics is that the experienced pleasantness (EP) from consuming a good depends only on its intrinsic properties and on the state of the individual (1). Thus, the pleasure derived from consuming a soda should depend only on the molecular composition of the drink and the level of thirst of the individual. In opposition to this view, a sizable number of marketing actions attempt to influence EP by changing properties of commodities, such as prices, that are unrelated to their intrinsic qualities or to the consumer's state. This type of influence is valuable for companies, because EP serves as a learning signal that is used by the brain to guide future choices. For example, when facing the choice between previously experienced restaurants, one would tend to avoid locales where previously meals were unsavory. Contrary to the basic assumptions of economics, several studies have provided behavioral evidence that marketing actions can successfully affect EP by manipulating nonintrinsic attributes of goods. For example, knowledge of a beer's ingredients and brand can affect reported taste quality (2, 3), and the reported enjoyment of a film is influenced by expectations about its quality (4). Even more intriguingly, changing the price at which an energy drink is purchased can influence the ability to solve puzzles (5).

Despite the importance and pervasiveness of various marketing actions, very little is known about the neural mechanisms through which they affect decisions made by individuals. An exception is a previous study demonstrating that knowledge of the brand of a culturally familiar drink, such as Coke, increases activation in the hippocampus, parahippocampus, midbrain, dorsolateral prefrontal cortex, and thalamus (6). The authors of the previous study interpreted such activity as evidence for retrieval of brand information during the consumption experience.

Here, we propose a mechanism through which marketing actions can affect decision making. We hypothesized that changes in the price of a product can influence neural computations associated with EP. This hypothesis is based on previous findings showing that affective expectations influence appraisals made about hedonic

experiences and, through this, the actual quality of experiences (2, 7, 8). Consider, for example, the experience of an individual sampling a wine for which he or she has information about its retail price. Because perceptions of quality are known to be positively correlated with price (9), the individual is likely to believe that a more expensive wine will probably taste better. Our hypothesis goes beyond this by stipulating that higher taste expectations would lead to higher activity in the medial orbitofrontal cortex (mOFC), an area of the brain that is widely thought to encode for actual experienced pleasantness (6, 10–16). The results described below are consistent with this hypothesis. We found that the reported price of wines markedly affected reported EP and, more importantly, also modulated the blood-oxygen-level-dependent (BOLD) signal in mOFC.

To investigate the impact of price on the neural computations associated with EP, we scanned human subjects ($n = 20$) using fMRI while they sampled different wines and an affectively neutral control solution, which consisted of the main ionic components of human saliva (17). We chose wine as a stimulus, because it is relatively easy to administer inside the scanner using computerized pumps, it induces a pleasurable flavor sensation in most subjects, and it varies widely in quality and retail price. Subjects were told they were sampling five different Cabernet Sauvignons, that the purpose of the experiment was to study the effect of degustation time on perceived flavors, and that the different wines would be identified by their retail prices (see Fig. 1A). Unbeknown to the subjects, the critical manipulation was that there were only three different wines, and two of them (wines 1 and 2) were administered twice, one identified at a high price and one at a low price. For example, wine 2 was presented half of the time at \$90, its retail price, and half of the time at \$10. Thus, the task consisted of six trial types: \$5 wine (wine 1), \$10 wine (wine 2), \$35 wine (wine 3), \$45 wine (wine 1), \$90 wine (wine 2), and neutral solution. The wines were administered in random order, simultaneously with the appearance of the price cue. Subjects were asked to focus on the flavor of the wine during the degustation period and entered taste pleasantness or taste intensity ratings in every other trial (Fig. 1B).

Results

Modulation of Reported Pleasantness and Taste Intensity by Price. We measured the impact of price information on EP by comparing the mean reported liking rating for wines 1 and 2 when administered at a high vs. a low price. We found significant differences for both wines ($P < 0.001$, Fig. 1C). In addition, reported pleasantness was correlated with wine prices ($r = 0.59$, $P < 0.000$). We could not find a similar behavioral effect for intensity ratings (Fig. 1D). To explore further the role of prices on

Author contributions: H.P., J.O., B.S., and A.R. designed research; H.P. performed research; H.P. analyzed data; and H.P., J.O., B.S., and A.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

*To whom correspondence should be addressed. E-mail: rangel@hss.caltech.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0706929105/DC1.

© 2008 by The National Academy of Sciences of the USA

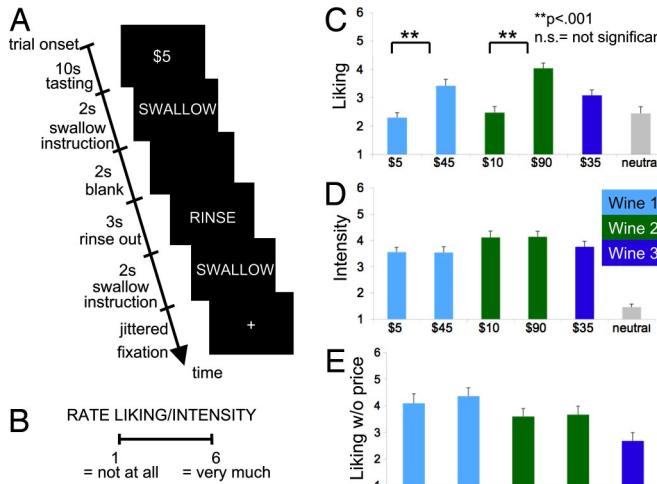


Fig. 1. Experimental design and behavioral results. (A) Time course for a typical trial. (B) Reported pleasantness and intensity rating scales. (C) Reported pleasantness for the wines during the cued price trials. (D) Taste intensity ratings for the wines during the cued price trials. (E) Reported pleasantness for the wines obtained during a postexperimental session without price cues.

experienced pleasantness, we administered a follow-up behavioral session 8 weeks after the main experiment, during which wines were presented without price information. As expected, in this case, there were no reported differences among the wines (Fig. 1E). Interestingly, while the pleasantness ratings were increasing in price in the main experimental task (Fig. 1C), they were not in the postscanning blind test (Fig. 1D). A potential concern with these behavioral results is they might exhibit “experimenter demand” effects. In particular, some subjects might deem it inappropriate to report to the experimenter that a cheaper wine tastes better.

Modulation of Brain Correlates of Experienced Pleasantness by Price. We analyzed brain imaging data using two different general linear models [see *Materials and Methods* and *supporting information (SI Text)* for details]. First, we looked for brain areas whose activity increased with the price of wine. More concretely, we estimated the BOLD response to each of the liquids at degustation and swallowing and then analyzed the contrasts “high–low price” at degustation separately for wines 1 and 2. Fig. 2B and C describe the results of this contrast for wine 1. We found increased activation in the left mOFC and the left ventromedial prefrontal cortex (vmPFC). Another cluster was found in a superior part of the vmPFC adjoining the rostral anterior cingulate cortex (rACC). We also found increased activation in the dorsolateral prefrontal cortex, visual cortex, middle temporal gyrus, and cingulate gyrus (see SI Table 2, upper). As shown in Fig. 2E and F, the contrast generated similar results for wine 2; increased activation was observed in the bilateral mOFC, vmPFC, and rACC. In addition, for the wine 2, we also found activation changes in the amygdala, lateral parts of the OFC, dorsolateral prefrontal cortex, inferior and middle temporal gyrus, and posterior cingulate cortex (see SI Table 3, upper).

A comparison of SI Tables 2, upper, and 3, lower, or of the relevant figures, suggests there might be small differences on the areas of the medial prefrontal cortex activated by the two wines. SI Fig. 5 shows that, as the statistical threshold is lowered, these differences disappear. To investigate this further, we performed a conjunction analysis to identify areas in which brain activity was higher on the high price condition for both wines. As shown in Fig. 3A, bilateral mOFC and adjoining rACC exhibited this pattern.

To investigate whether the increase in price had a differential effect on the two wines, we performed an interaction analysis

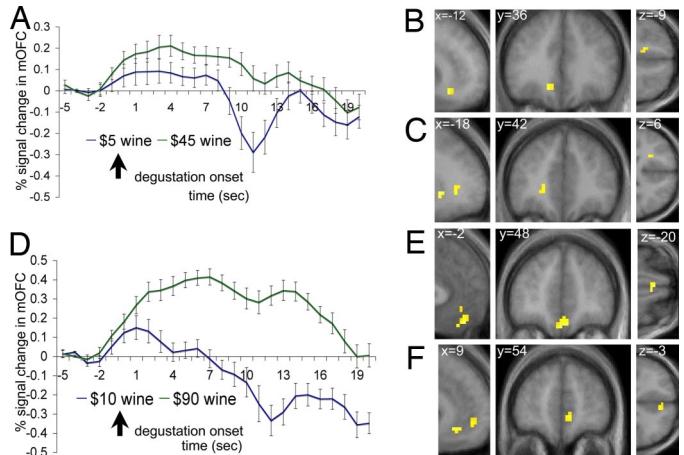


Fig. 2. The effect of price on each wine. (A) Wine 1: averaged time courses in the medial OFC voxels shown in B (error bars denote standard errors). (B) Wine 1: activity in the mOFC was higher for the high- (\$45) than the low-price condition (\$5). Activation maps are shown at a threshold of $P < 0.001$ uncorrected and with an extend threshold of five voxels. (C) Wine 1: activity in the vmPFC was also selected by the same contrast. (D) Wine 2: averaged time courses in the medial OFC voxels shown in E. (E) Wine 2: activity in the mOFC was higher for the high- (\$90) than for the low-price condition (\$10). (F) Wine 2: activity in the vmPFC was higher for the same contrast.

(see *Materials and Methods* and *SI Text* for details). We found that the effect of price on mOFC activity was higher for the cheap \$5 wine than for the expensive \$90 wine. This suggests that the effect of a price increase on mOFC activity might be larger at low than at high prices.

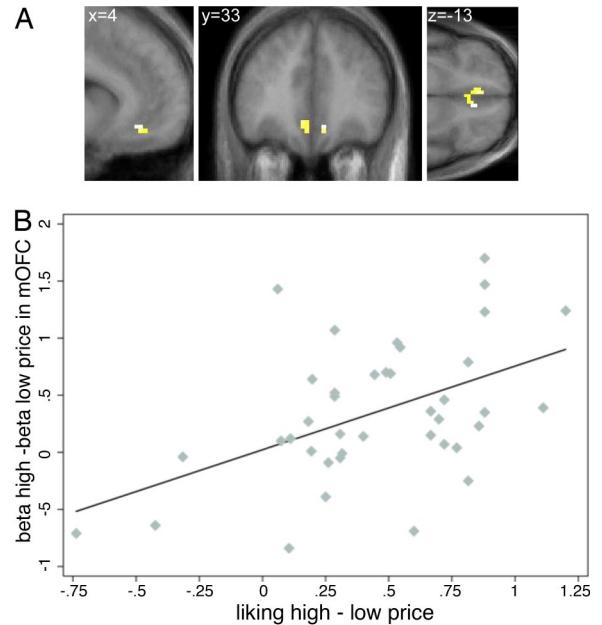


Fig. 3. The effect of price on both wines. (A) Conjunction analysis. Activity in the mOFC/rACC was higher in the high- than in the low-price condition for both wines 1 and 2. (B) Correlation of behavioral and BOLD responses ($r = 0.49$, $P < 0.001$). Each point denotes an individual wine pair. The horizontal axis measures the change in reported pleasantness between the high- and low-price conditions. The vertical axis computes an analogous measure using the betas from the general linear model in a 5-mm spherical volume surrounding the area depicted in A.

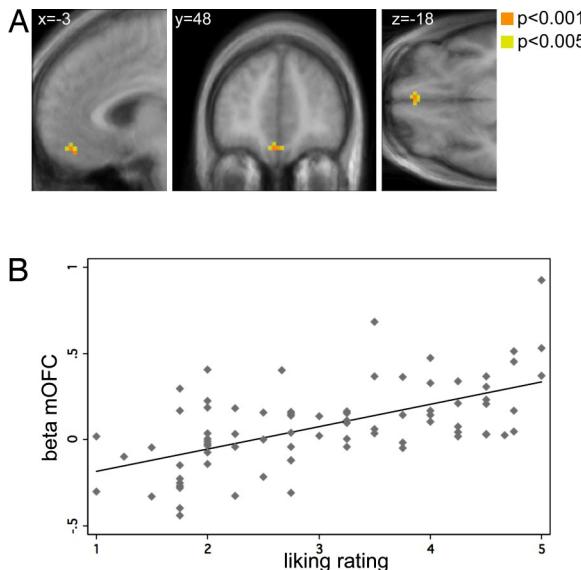


Fig. 4. Neural correlates of liking ratings. (A) Activity in the mOFC and the midbrain correlated with the reported pleasantness of the six liquids at degustation time. For illustration purposes, the contrast is shown both at $P < 0.001$ and $P < 0.005$ uncorrected and with an extend threshold of five voxels. (B) Correlation of pleasantness ratings and BOLD responses ($r = 0.593$, $P < 0.000$). Each point denotes a subject-price pair. The horizontal axis measures the reported pleasantness. The vertical axis computes the betas from the general linear model in a 5-mm spherical volume surrounding the area depicted in A.

Discussion

The main hypothesis of this study was that an increase in the perceived price of a wine should, through an increase in taste expectations, increase activity in the mOFC. The results described above provide evidence consistent with the hypothesis. The hypothesis was motivated by several previous studies, which have shown that activity in the mOFC is correlated with behavioral pleasantness ratings for odors (10–13), tastes (6, 14, and 15), and even music (16). This, together with our behavioral results and the additional imaging results described below, support the interpretation that, by modulating the activity in the mOFC, changes in the price of a wine might lead to a change in the actual EP derived from its consumption.

We performed two additional analyses to provide further support for this interpretation. First, for each individual and wine, we computed the change in reported EP between the high and low price conditions. We also computed the analogous difference in parameter estimates for the BOLD response from the general linear model in an area surrounding the mOFC. Fig. 3B shows that the neural and behavioral estimates were positively and highly correlated ($r = 0.49$, $P < 0.001$). Second, we verified that the results of the previous literature also held in our study by estimating a different general linear model and looking for brain regions whose activity was correlated with reported EP from sampling the different stimuli (see *SI Text* for details). The results replicated the findings of previous studies: activity in the mOFC was correlated with absolute reports of pleasantness (Fig. 4).

Importantly, we did not find evidence for an effect of prices on areas of the primary taste areas such as the insula cortex, the ventroposterior medial nucleus of the thalamus, or the prabial nuclei of the pons. A natural interpretation is that the top-down cognitive processes that encode the flavor expectancies are integrated with the bottom-up sensory components of the wine in the mOFC, thus modulating the hedonic experience of flavor, but that the flavor expectancies generated by the change in prices do not impact more basic sensory representa-

tions. Interestingly, an analogous mechanism has been proposed for pain placebo effects (7).

Our results have implications for several disciplines. First, the EP signal plays a central role in neuroeconomics, because it serves as a teaching signal that guides future behavior. Unfortunately, very little is known about the factors that affect the neural computation of this signal. A natural starting hypothesis is the economic view, which states that EP depends only on the sensory properties of the item being consumed (i.e., its molecular properties) and the state of the consumer. Our results suggest that the brain might compute EP in a much more sophisticated manner that involves integrating the actual sensory properties of the substance being consumed with the expectations about how good it should be. It is important to emphasize that it might be adaptive for the brain to do this. To make good decisions in the future, the brain needs to carry out good measurements of the quality of current experiences. In a world of noisy measurements, the use of prior knowledge about the quality of an experience provides additional valuable information. A related study (13) provides additional supporting evidence for this point by showing that giving a cognitive label to an ambiguous odor (“cheddar cheese” or “body odor”) can affect both subjective pleasantness reports and neural activity related to EP. Unlike the current paper, however, de Araujo *et al.* (13) do not provide evidence that marketing actions, such as pricing, can affect neural correlates of EP.

Second, our findings also have implications for marketing. Whereas there is ample behavioral evidence that various marketing actions are successful in influencing the EP of individuals, that they can modulate neural representations of this signal had not been reported before. Furthermore, the neural findings also provide some clues about the mechanisms involved. In particular, it seems that price changes modulate the representations of experienced utility but not the encoding of the sensory properties of taste in the primary gustatory cortex.

Third, our results have implications for economics. EP is an important component of experienced utility, which is the economist’s term for subjective well being. We show that, contrary to the standard economic view, EP depends on nonintrinsic properties of products, such as the price at which they are sold. It then follows that marketing manipulations might affect subjective perceptions of well being. This raises several difficult questions for the field. Should the effect of prices on experienced utility be counted as real economic well being or as a mistake made by individuals? To what extent are measurable differences in preferences based on intrinsic differences between products and price effects we have identified? What happens to the efficiency of competitive markets when firms can influence experienced utility by changing the price of items?

An important task for future research is to develop a more complete characterization of the range of marketing actions that can influence the neural computation of EP. We conjecture that any action affecting expectations of product quality, such as expert quality ratings; peer reviews; information about country of origin, store, and brand names (especially those associated with luxury products); and repeated exposure to advertisements might lead to effects similar to those identified here.

Materials and Methods

Subjects. Twenty normal-weight subjects participated in the experiment (11 males, ages 21–30; mean age, 24.5 yr). One additional subject participated in the experiment but was excluded from the analysis, because he reported being confused about the task during a debriefing at the end of the experiment. All subjects were right-handed and healthy; had normal or corrected-to-normal vision and no history of alcohol abuse, psychiatric diagnoses, or neurological or metabolic illnesses; and were not taking any medications that interfere with the performance of fMRI. All subjects were screened for liking, and at least occasionally drinking, red wine. At the beginning of each experiment, subjects were required to show an official form of identification to

provide evidence they were >21 yr old. Subjects were informed about the experiment and gave written consent before participating. California Institute of Technology's institutional review board approved the study.

Stimuli. During the course of the fMRI experiment, subjects sampled three different Cabernet Sauvignon wines and an affectively neutral tasteless control solution that consisted of the main ionic components of human saliva (25 mM KCl and 2.5 mM NaHCO₃). The wines were administered in random order and simultaneously with the appearance of a price identifier. Two of the three wines were administered twice, once identified by their actual retail price and once by a 900% markup (wine 1: \$5 real retail price, \$45 fictitious price) or a 900% reduction (wine 2: \$90 real retail price, \$10 fictitious price). The third wine was used as a distracter and was identified by its retail price (wine 3: \$35). We also carried out a follow-up behavioral tasting session 8 weeks after the main experiment. In this postexperimental session, the wines were presented without price information (see *SI Table 1* for a complete description).

In each trial, 1 ml of each stimulus was delivered by a system of electronic syringe pumps (one for each stimulus) positioned in the scanning control room. These pumps transferred the stimuli to the subjects via ≈10-m polyethylene plastic tubes (6.4-mm diameter) and a perfusion manifold. The perfusion manifold allowed six incoming tubes to be connected to one output tube with a minimum of dead space to avoid the mixing of the wines. The subjects were instructed to hold the output tube between their lips like a straw while they lay in a supine position in the scanner. We made an effort to keep the tubes free of air bubbles to avoid wine oxygenation. Between experiments, the wines were preserved with a special corking system, which is typically used by wineries and wine shops to prevent oxidation.

The stimulus presentation and response recording was controlled by Co-Genet 2000 (Wellcome Department of Imaging Neuroscience) installed on a computer positioned in the control room that also received trigger pulses from the scanner to control timing. The visual stimuli were presented by using video goggles (Resonance Technologies).

Task. The task consisted of six trial types: \$5, \$10, \$35, \$45, and \$90 wines and a neutral liquid. Each trial type occurred 16 times during the course of the experiment, resulting on a total of 96 trials. Subjects were instructed to sample the liquid on each trial while it was on their mouth (for a period of 10 s), to evaluate its pleasantness during this time, and to swallow only when instructed. Between every wine administration, there was a rinse period in which the neutral solution was delivered. The rinse period was implemented to avoid taste spillovers across trials. Trials were separated by a random intertrial interval drawn from a Poisson distribution with a mean of 10 s (see Fig. 1A for a detailed description of the timing of each trial).

In every other trial, subjects were instructed to enter a rating of either flavor pleasantness or taste intensity. Thus, a total of four pleasantness ratings and four taste intensity ratings were sampled for each liquid. We used a six-point rating scale (1 = do not like it at all/not intense at all; 6 = like it very much/very intense) (see Fig. 1B). The timing of rating trials was identical to nonrating trials, except that, after swallowing the liquid and before rinsing, subjects were given 6 s to enter their ratings.

After an initial instruction period, subjects were trained on the use of the response boxes to enter the ratings. Given the large number of potential ratings, subjects entered ratings with both hands (three ratings for each hand). The assignment of ratings to buttons was counterbalanced across subjects to avoid motor artifacts.

Unbeknown to the subjects, the critical manipulation was that the \$5 and \$45 wines and the \$10 and \$90 wines were identical. This manipulation was not revealed to the subjects. Instead, the subjects were told they would be sampling five different Cabernet Sauvignons, that the purpose of the experiment was to study the effect of sampling time on perceived flavors, and that the different wines will be identified by their retail prices. Evidence for the success of our cover story was that all subjects reported at the end of the experiment being able to taste five different wines. Although the experiment contains an element of deception, subjects were not debriefed after the experiment to avoid contaminating California Institute of Technology's small subject pool.

Data Acquisition and Preprocessing. The brain imaging was conducted in a 3-T Siemens Trio MRI scanner (Siemens). We acquired gradient echo T₂ weighted echoplanar images (EPI) with BOLD contrast and used a special

sequence designed to optimize functional sensitivity in orbitofrontal cortex (18). This consisted of a tiled acquisition in an oblique orientation at 30° to the AC-PC line. In addition, we used an eight-channel phased array coil that yields a 40% signal increase in OFC over the standard head coil. The sequence enabled 32 axial slices of 3-mm thickness and 3 mm in-plane resolution that could be acquired with a TR of 2 s. A T1-weighted structural image was also acquired for each subject. Functional imaging data were acquired in four separate sessions of ≈13 min each.

To detect transient head movements due to swallowing, we attached a 1.5-cm-long copper coil with a radius of 0.5 cm to the neck of each subject. The setup was similar to those used by previous studies in which liquid food had been administered (14). Small movements of the coil induced a current in the magnetic field that could be detected after amplification using an EEG system positioned in the scanner room (Biopac Systems). This produced a time series of events reflecting transient larynx movement that was used in the general linear model (GLM) in two ways. First, the time series of the signal detected by the coil was added as an additional motion regressor in the GLM as regressor of no interest; this was done to take out the variance due to swallowing induced head movement. Second, swallowing events were either assigned to experimental conditions or classified as noninstructed swallowing. The former were used to correct the timing of swallowing onsets for each stimulus type that were entered in the GLM as a regressor of interest (see *SI Text* for details about the models). The latter were entered into the data analysis as a regressor of no interest.

fMRI data analysis was performed by using the Statistical Parametric Mapping software (SPM05; Wellcome Department of Imaging Neuroscience). We applied the following preprocessing steps to the imagining data: (i) slice-timing correction (centered at TR/2) (ii) realignment to the last volume, (iii) spatial normalization to a standard T2* template with a resampled voxel size of 3 mm³, (iv) spatial smoothing using a Gaussian kernel with full width at half maximum of 8 mm, and (5) intensity normalization and high-pass temporal filtering (filter width 128 s). The structural T1 images were coregistered to the mean functional EPI images for each subject and normalized using parameters derived from the EPI images.

fMRI Data Analysis 1: Influence of Price on Wine Sampling. We estimated a general linear model in which the delivery of each of the six different liquids was entered as a regressor of interest. In addition, the swallows of each liquid type were also entered as separate regressors. Each of these regressors plus additional regressors of no interests were convolved with a canonical hemodynamic response function. We then calculated first-level single-subject contrasts to compare the administration of an identical wine at a high minus a low price. Finally, for each of these first-level contrasts, we calculated a second-level group contrast using a one-sample *t* test (see *SI Text* for details).

We also performed an interaction analysis to compare the effect of prices on the two wines. We calculated a first-level single-subject contrast to compare the administration of the low-quality wine at a high minus a low price minus the administration of the high quality wine at a high minus a low price. We also calculated a first-level single-subject contrast to compare the administration of the high-quality wine at a high minus a low price minus the administration of the low-quality wine at a high minus a low price. For each of these first-level contrasts, we calculated a second-level group contrast using a one-sample *t* test (see *SI Text* for details).

fMRI Data Analysis 2: Neural Representation of Experienced Pleasantness. We estimated a second general linear model in which the delivery and swallow of all liquids (independently of type) were entered as regressors. In addition, we entered the reported pleasantness and intensity ratings as parametric modulators for both regressors. We then computed first-level single-subject contrasts for the parametric modulators at both liquid sampling and swallowing by reported EP and taste-intensity ratings. Finally, for each of these first-level contrasts, we calculated a second-level group contrast using a one-sample *t* test (see *SI Text* for details).

ACKNOWLEDGMENTS. We thank Vivian Valentin, Jan Glaescher, and Axel Linder for their help with this project and Hackjin Kim, Sam Huang, and Shawn Wagner for developing the coil we used to detect swallowing movement and for providing support to process the swallowing signal. Financial support from the National Science Foundation (Grant SES-0134618) is gratefully acknowledged. This work was also supported by a grant from the Gordon and Betty Moore Foundation to the California Institute of Technology Brain Imaging Center.

1. Kahneman D, Wakker PP, Sarin R (1997) Back to Bentham? Explorations of Experienced Utility. *Q J Econ* 112:375–405.
2. Lee L, Frederick S, Ariely D (2006) Try it, you'll like it: The influence of expectation, consumption, and revelation on preferences for beer. *Psychol Sci* 17:1054–1058.
3. Allison RJ, Uhl KP (1964) Influence of Beer Brand Identification on Taste Perception. *J Market Res* 1:36–39.
4. Klaaren KJ, Hodges SD, Wilson TD (1994) The role of affective expectations in subjective experience and decision-making. *Soc Cognit* 12:77–101.

5. Shiv B, Carmon Z, Ariely D (2005) Placebo Effects of Marketing Actions: Consumers May Get What They Pay For. *J Market Res* XLII:383–393.
6. McClure SM, et al. (2004) Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44:379–387.
7. Wager T (2005) The neural bases of placebo effects in pain. *Curr Dir Psychol Sci* 14:175–179.
8. Wilson TD, Klaaren KJ (1992) in *Review of Personality and Social Psychology*, ed Clark MS (Sage, Newbury Park, CA), Vol 14, pp 1–31.
9. Rao A, Monroe KB (1989) The effect of price, brand name, and store name on buyers' perceptions of product quality: An integrative review. *J Market Res* 26:351–357.
10. de Araujo IE, Kruegelbach ML, Rolls ET, McGlone F (2003) Human cortical responses to water in the mouth, and the effects of thirst. *J Neurophysiol* 90:1865–1876.
11. Anderson AK, et al. (2003) Dissociated neural representations of intensity and valence in human olfaction. *Nat Neurosci* 6:196–202.
12. de Araujo IE, Rolls ET, Kruegelbach ML, McGlone F, Phillips N (2003) Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *Eur J Neurosci* 18:2059–2068.
13. de Araujo IE, Rolls ET, Velazco MI, Margot C, Cayeux I (2005) Cognitive modulation of olfactory processing. *Neuron* 46:671–679.
14. Kruegelbach ML, O'Doherty J, Rolls ET, Andrews C (2003) Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb Cortex* 13:1064–1071.
15. Small DM, et al. (2003) Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39:701–711.
16. Blood AJ, Zatorre RJ (2001) Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci USA* 98:11818–11823.
17. Francis S, et al. (1999) The representation of pleasant touch in the brain and its relationship with taste and olfactory areas. *NeuroReport* 10:453–459.
18. Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage* 19:430–441.



ELSEVIER

Brain Research Bulletin 67 (2005) 428–437

BRAIN
RESEARCH
BULLETIN

www.elsevier.com/locate/brainresbull

Arousal and consumer in-store behavior

Andrea Groeppel-Klein *

*Chair of International Marketing, Consumer Behavior, and Retailing at the European University Viadrina,
Grosse Scharrnstrasse 59, 15230 Frankfurt, Oder, Germany*

Available online 18 July 2005

Abstract

From a psychophysiological point of view, arousal is a fundamental feature of behavior. As reported in different empirical studies based on insights from theories of consumer behavior, store atmosphere should evoke phasic arousal reactions to attract consumers. Most of these empirical investigations used verbal scales to measure consumers' perceived phasic arousal at the point-of-sale (POS). However, the validity of verbal arousal measurement is questioned; self-reporting methods only allow a time-lagged measurement. Furthermore, the selection of inappropriate items to represent perceived arousal is criticized, and verbal reports require some form of cognitive evaluation of perceived arousal by the individual, who might (in a non-measurement condition) not even be aware of the arousal. By contrast, phasic electrodermal reaction (EDR) has proven to be the most appropriate and valid indicator for measuring arousal [W. Boucsein, Physiologische Grundlagen und Meßmethoden der dermalen Aktivität. In: F. Rösler (Ed.), Enzyklopädie der Psychologie, Bereich Psychophysiologie, Band 1: Grundlagen und Methoden der Psychophysiologie, Kapitel, Vol. 7, Hogrefe, Göttingen, 2001, pp. 551–623] that could be relevant to behavior. EDR can be recorded simultaneously to the perception of stimuli. Furthermore, telemetric online device can be used, which enables physiological arousal measurement while participants can move freely through the store and perform the assigned task in the experiments. The present paper delivers insights on arousal theory and results from empirical studies using EDR to measure arousal at the POS.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Electrodermal reactions; Buying behavior; In-store stimuli

1. Introduction

As reported in different empirical studies based on insights from theories of consumer decision-making and on insights from environmental psychology, store atmosphere must evoke phasic arousal reactions to attract consumers. Most of these empirical investigations used verbal scales to measure consumers' perceived arousal at the point-of-sale. However, the validity of verbal arousal measurement can be questioned. On the one hand, traditional interviews are mostly done after the shopping trip through a store or a particular part of the store so that customers have to remember their perceived arousal and emotions. Thus, self-reporting methods only allow a time-lagged measurement. On the other hand, the preferential use of verbal scales to measure psychophysiological responses is discussed controversially. By contrast,

electrodermal activity (EDA) is considered to be a valid and also very sensitive indicator that responds clearly to the smallest variation in arousal. Latest technical developments offer larger storage capacity enabling telemetric measurement of electrodermal activity, which in turn makes field experiments in retail stores possible. The present paper delivers insights on arousal theory and presents empirical results from studies using EDA to measure arousal at the point-of-sale.

2. A psychophysiological perspective: insights from arousal theory

From a psychophysiological point of view, arousal is a fundamental feature of behavior. It can be defined as the neurophysiological basis underlying all processes in the human organism. Thus, arousal is the basis of emotions, motivation, information processing, and behavioral reactions [2,3,10,27,32]). Arousal can vary from deep sleep through

* Tel.: +49 335 5534 2870; fax: +49 335 5534 2275.

E-mail address: groeppel@euv-frankfurt-o.de.

moderate stages of increase up to panic. Basically, a distinction can be made between tonic and phasic arousal. Tonic arousal refers to a relatively long-term state of consciousness that changes slowly due to long-lasting or extremely intensive stimuli. Phasic arousal arises in response to specific stimuli, resulting in short-term variations in the arousal level. It indicates a ready state of the body for reaction. It is closely related to attention, i.e. enhanced sensitivity of the organism to relevant stimuli and stimuli processing, while irrelevant stimuli are filtered and not processed. Both internal (thought and metabolic processes) and external stimuli can cause arousal [6,7]. Phasic arousal might be the driving force for decision-making processes and approach behavior (i.e. time and money spent in a store) at the point-of-sale (see Section 3).

Early unidimensional concepts of arousal theory describe the reticular activating system (RAS), which comprises the sensoric inflow, the reticular formation (RF), as well as cortical, hypothalamic, and thalamic areas, as the fundamental component for the formation of arousal [5,6]. The RF is a complex network of fibers and cell bodies in the core of the brain stem. It is involved in the filtering process of sensory information from the central nervous system (e.g. from visual, haptic, and acoustic stimulation). According to these early theoretical approaches in arousal research [13,29], all sensory and motor nerve fibers should enhance general arousal of the reticular formation via collaterals. In turn, the RF should indistinctly activate large parts of the central nervous system [6]. Thus, the unidimensional approach suggested a correlation between diverse physiological outcomes of arousal, such as heart rate, blood pressure, electroencephalogram (EEG), or electrodermal activity.

Since empirical findings were not consistent with this unidimensional approach, recent efforts yielded towards a

more complex arousal theory [5,6,16,28]. Boucsein [6], for example, presented a three-dimensional theoretical framework based on neurophysiological insights on arousal and information processing (see Fig. 1). This model has been confirmed by a PET-study by Fredrikson et al. [19]. The model also considers the dependency between arousal and emotion/motivation as well as the effects on central and peripheral psychophysiological parameters. The first dimension of this framework is in accordance with the concept of the previous unidimensional approach. That means that the RF forms the physiological basis of this dimension. EEG can serve as an indicator. Results from the first dimension are perceived as general activation with a vigilant feeling and an alert state of mind. The second dimension, the “affect–arousal” system, comprises primarily emotional components of arousal. Headed by the Amygdala, attention, orienting reflex, and overall behavior is enhanced via hypothalamic reactions. The physiological outcome results in phasic cardiovascular (heart rate) and/or tonic electrodermal variations. When it comes to behavior and perception, processes of this second dimension lead to defense reactions and negative emotions. Attention that turns into an orienting reflex may also directly impinge on the third system. This so-called “preparatory activation” system is basically encompassing motivational aspects of arousal. Expectations are transformed into a ready state for reaction. This part of the system interacts, especially, with motor and pre-motor activation of behavior, and with positive emotion. Phasic electrodermal amplitude, for example, can serve as an indicator.

For marketing purposes, and especially, when testing the impact of in-store stimuli on customers’ arousal and buying behavior, the “preparatory activation” system, and in part, the “affect–arousal” system (here, especially, attention and orienting reflex), are of major relevance [23].

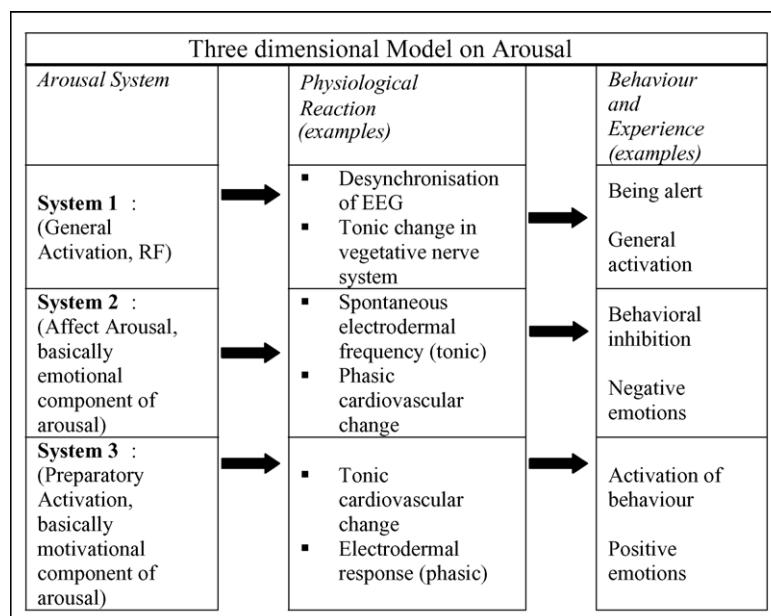


Fig. 1. Simplified own illustration of the three-dimensional arousal model of Boucsein and Aktivierung [6].

3. Arousal and in-store behavior

For more than 20 years, the empirical investigation of consumers' responses to store atmosphere and visual merchandising concepts has been an interesting research area and it has now become widely accepted that an appealing store atmosphere has a major impact on consumers' in-store behavior and store assessment [11,18,20–22,24,30,36,38].

In summary, the results clearly confirm that a store atmosphere that evokes pleasure, can increase consumers' spending, time spent in the store, and the desire to come back and to recommend the store to others. However, the role of perceived arousal at the point-of-sale is discussed controversially. Some of the cited studies show a significant positive influence of an arousal stimulating-atmosphere on store-assessment and in-store behavior (e.g. [11,23]), others do not support these findings or even show a converse relationship (e.g. [12]). One reason for these contradictory results may be that the studies used different methods to measure arousal, which questions the validity of the results (see Section 4). A second reason could be that not all investigated environmental settings really offered arousal-increasing stimuli. Three groups of arousal inducing stimuli can be distinguished [27]: (1) affective stimuli evoke pleasant or unpleasant emotions due to innate stimulus-response mechanisms, or preceding conditioning. Especially, key-stimuli, such as little children, erotic stimuli, and symbols from nature increase the arousal appealing mannequins, scents, warm colors or plants are possible affective in-store stimuli. (2) Intense stimuli produce an effect because of their physical properties. As salient information, such stimuli automatically engender orienting responses. At the POS intense stimuli could be flashy colors, spotlights, visually, striking price-tags, or some retailers use huge flat screens presenting movies, music, etc. (3) Collative stimuli are novel, startling, or different stimuli arrangements (e.g. vivid-looking mannequins or unexpected decorations).

Furthermore, little contribution can be found in the literature neither on the impact of store atmosphere on buying decisions nor on the impact of arousal on decision processes evoked by the store atmosphere. Therefore, in our studies, we do not only analyze the influence of store atmosphere and visual merchandising concepts on arousal perceived at the point-of-sale (e.g. [23]), but we are also interested in comparing phasic arousal reactions of buyers with those of non-buyers [25].

The relevance of taking non-buyers into consideration, results from diverse studies reporting that shopping trips can have a hedonic value irrespective of whether or not something is purchased [1,23,26] and that consumer behavior does not always satisfy functional or economic needs [17,34]. According to Babin et al. [1], hedonic shopping results "more from fun and playfulness than from task completion (...)" and "reflects shopping's potential entertainment and emotional worth (...); increased arousal, heightened involvement, perceived freedom, fantasy fulfillment, and escapism may all indicate a hedonically valuable shopping experience".

On the one hand, of course, appealing store designs can evoke positive feelings at the POS for both, buyers and non-buyers. Since we regard arousal as the physiological basis of emotions, we can reasonably conclude that a positive store atmosphere will also increase consumer arousal independently of any purchase. On the other hand, we know that highly appealing stores designs also provoke more purchases (e.g. [11,20]). This leads to the notion that "hedonic buyers" will reveal even higher arousal than "hedonic non-buying browsers". In addition, according to the previous research cited earlier, making a real decision implies more effort than leaving it, which is another reason for higher phasic arousal reactions of buyers. During economic downturns, many retailers complain about the tendency of consumers to avoid stores and often to refrain altogether from buying. Due to this "consumption lethargy", non-buyers might be less activated irrespective of the store ambience. This argument lends further support to the thesis that on average, buyers are more aroused than non-buyers. Thus, this paper focuses on two research questions:

- (1) Do different visual merchandising concepts evoke different phasic arousal responses of consumers, and can arousal be related to emotions, and in-store behavior?
- (2) Are buyers more aroused at the point-of-sale than non-buyers, and can different arousal patterns be detected?

4. Measurement of arousal

Arousal is definitely considered as being fundamental to behavior that is related to emotions and behavior [3,8] and as previously shown, arousal plays a major role for customers' interaction with the store environment. Yet, despite diverse methods applied in order to measure arousal, the variable is extremely difficult to indicate validly. Among the applied methods are verbal scales like the PAD scale [33], non-verbal scales (e.g. color-pattern scale [4,31], and self-reporting methods.

The color and pattern scale contains more or less intense colors or complex patterns. Test persons have to select from the scale one color or a pattern that reflects their subjectively perceived arousal level. Meyer-Hentschel [31] assumes that excited consumers choose instinctively different colors or patterns than less aroused test persons. However, the color-pattern scale has not yet been implemented often enough to support valid conclusions. Furthermore, irrespective of their arousal reaction, test persons may have an enduring preference for special colors or patterns. Responses from self-reporting methods often suffer from biases [40]. Another way to investigate arousal at the point-of-sale – but in fact, rather measuring emotion – is the observation of non-verbal communication, like orienting reflexes, or especially, facial features, e.g. via facial affect scoring technique (FAST) [15] or facial acting coding system (FACS) [14]. Today, also computer-aided face-recognition methods are used, e.g.

Zlochower and co-workers [9,37]. Furthermore – as mentioned before – a variety of different verbal methods has been used to measure perceived arousal at the POS. However, the results derived from these scales are controversial, and their discriminant validity has been criticized [38,40]. Basically, there are three reasons for these criticisms: (1) selection of appropriate items to represent perceived arousal; (2) verbal reports require some form of cognitive evaluation of perceived arousal by the individual, who might (in a non-measurement condition) not even be aware of the arousal; (3) the fact that verbal measures are taken some time later than the actual experience of the situation [23].

In contrast to these methods, psychophysiological measures, such as heart rate, EEG, and EDA, or studies with brain-imaging technologies (e.g. fMRI or PET) offer most valid indicators. However, at present neither EEG nor brain-imaging technology can be used in field experiments.

Electrodermal activity in general “is regarded as a sensitive and valid indicator for the lower arousal range, reflecting small, mostly cognitively conditioned variations in arousal” [5]. Willingly influencing test results is almost impossible. Contrary to the heart rate, the electrodermal reactivity (EDR) points to even the very smallest psychological change [5]. It is thus delivering the most sensitive indicator of arousal that might potentially be relevant to behavior. Furthermore, as we will see later, EDR can also be measured at the point-of-sale. EDR can, therefore, deliver a valuable contribution to further research on consumers’ shopping behavior induced by arousal.

5. Electrodermal activity as psychophysiological indicator for arousal

The human skin basically consists of two layers, the dermis, and the epidermis. The epidermis is located at the surface and consists of epithelial tissue. This layer is more callous, the closer to the surface. The comparatively thicker and deeper lying dermis consists of taut, fibrous connective tissue. Underneath the dermis, lies the subcutis, which contains the secretory part of the sweat glands, fatty tissue, and vessels that supply the body surface [5].

The epidermis is of great importance to EDA. Although it becomes dryer towards its outside layer as the regularly arranged cells become less tightly packed, there is a permanent insensible perspiration from the dermis via the epidermis even while no sweat gland activity occurs. This hydration depends on external and internal factors and leads to good electric conductivity of the skin, thus, making it possible to measure it by means of two electrodes attached to the skin. The conductivity of the skin is then transmitted to the computer by means of an amplifier. The conductivity of the skin is primarily responsible for tonic EDA, while active membrane processes following upon a nerve impulse bring about phasic EDA which turns into electrodermal reactivity [5]. Following the three-dimensional model of Boucsein [6], phasic electro-

dermal reactivity is the most crucial indicator for research concerning arousal at the point-of-sale.

Basically, EDA can be measured via endosomatic or exosomatic recording. While endosomatic recording does not use any external voltage to measure potential differences on the skin, exosomatic recording uses either direct current (dc) or alternating current (ac). Exosomatic recording with direct current is most frequently used to measure either skin conductance (SC), measured in microsiemens (μ S). Siemens is the unit to measure conductance between two objects or skin resistance (SR), measured in kilohm ($k\Omega$) (ohm is the unit to measure resistance between two objects). Each of these methods measures both, skin conductance/resistance level (tonic arousal) as well as skin conductance/resistance reactivity (phasic arousal) from the EDA-signal [5]. For practicality and safety reasons for test subjects, exosomatic recording must strongly be recommended for in-store investigations since endosomatic recording requires one of the electrodes being placed in an “inactive” site on the skin that has been pretreated (so-called “skin-drilling”). For exosomatic measurement, both electrodes are simply attached to the left palm of right-handed test persons, or vice versa for left-handers.

6. Measurement of EDR at the point-of-sale

In order to investigate the research questions arousal was measured by recording EDR, while the test subjects performed the assigned task in the experiments. In all studies, a telemetric online device was used, which enables physiological arousal measurement while participants can move freely through the store. Rapid movements, pressure on the electrodes, change in intimacy between electrodes and skin, as well as speech activity, may lead to artifacts resulting in specific patterns in the EDR curve. Such artifacts need to be excluded from the data [5,23]. However, due to new EDR-technology, “normal walking” does not lead to artifacts, therefore, we can use this method in field experiments. For registration, an exosomatic approach was chosen applying dc (0.4 V) and measuring skin conductance. The technical equipment runs with a 12-bit analog-to-digital (A/D) converter. Two Ag/AgCl electrodes were used and filled with a 0.5% NaCl electrode creme.

The following EDR parameters are of particular relevance to all experiments. The amplitude describes the strength of each phasic arousal reaction [5]. In order to calculate amplitudes, for all studies, we chose a minimum amplitude criterion of 0.05 μ S so as to exclude recording artifacts from the signal-to-noise ratio [5,39]. While registering EDR, overlapping amplitudes may occur, i.e. a second amplitude follows a first one, although the original baseline has not yet been reached. For our data analysis, overlapping amplitudes were evaluated, each by means of its own baseline, regardless of the recovery time of the preceding amplitude [5]. According to Steiger [35], the intensity of perceived arousal over a certain period of time, can be received by summing up all single ampli-

tudes so as to obtain the total amplitude which is the most important phasic arousal parameter in field experiments. The second important parameter is the so-called “frequency” of responses, which subsumes up all single reactions. Each skin conductance response demonstrates the particular attention of the individual towards an object in its environment [35].

7. Empirical studies

7.1. Studies 1 and 2

The first empirical study investigates the influence of different visual merchandising concepts on arousal, emotions, and buying behavior (for details, see Ref. [23]). The investigation was conducted in the fruit and vegetable department of two Austrian grocery stores. Both stores belong to the same retail chain and are located in a distance of a 10 min walk from another. The management of the retail chain pursues different marketing strategies with both stores. For “Store 1” (experimental store), insights from environmental psychology have been considered. Here, fruits and vegetables are presented in large-scale on broad and deep carriers and are arranged according to colors. Exotic fruits, fresh herbs, and flowers are used as eye-catchers (“affective stimuli”). Parts of the products are presented on sort of an “island” in the middle. The whole store is spacious and high-ceilinged, and is brightly illuminated. In “Store 2” (control store), the same assortment of fruits and vegetables is offered. Quality and prices are exactly identical. However, insights from environmental psychology have not been considered for the store decoration. The light appears less bright. Shelves are put alongside the wall, and display tables are arranged in a row in the middle of a long and narrow aisle. Contrary to “Store 1”, an extraordinary product presentation (for example, by an arrangement of the products according to colors) was not applied.

It can be assumed that “Store 1” which is decorated according to the principles of environmental psychology would evoke higher arousal than the ordinary “Store 2”. In the experimental store, 15 test persons were asked to do their usual shopping in the vegetable department, while EDR was registered. In the control store, we collected EDR data from 12 test persons. Test customers were asked to participate right after entering the store. Both, the EDR apparel and the telemetric device, were either laid in the shopping cart, or put in the test

Table 2
Correlation between emotion “joy” and arousal

Emotion	Total amplitude	Frequency
Joy (<i>n</i> = 27)		
Pearson correlation (two-tailed significance)	.464 (.015)	.420 (.029)

subject’s pocket. The data were transmitted via the telemetric device to the computer that was installed at the end of the fruit and vegetable department for online registration. From that part of the store, we also observed the participants. In both stores, the fruit and vegetable section is the first part of the store the test subjects walked through. Before they went on to another part of the store, electrodes were detached, and registration stopped. Each individual was then asked to answer a standardized questionnaire. The questionnaire was the same for both stores.

In order to gain valid parameters, we first excluded all artifacts from the recorded curves. Furthermore, the first part of the curves were not rated up to the point when the test subjects were used to the electrodes and their curve showed a consistent level (individual baseline = 0). When summing up the amplitudes of the test persons generated over their controlled shopping time, we found significant difference in total amplitudes for the two stores by means of a Mann–Whitney *U*-test (see Table 1). A Mann–Whitney *U* was chosen to replace independent samples *T*-tests since small samples do not satisfy the normal distribution assumption. This means that the experimental store evokes higher arousal than the control store. Furthermore, data show significantly higher frequency in the response rate of test persons in the experimental store compared to those in the control store. This indicates that the environment of the experimental store provides more information that has the potential to attract enough attention to be cognitively processed.

EDA measures arousal, however, the perceived emotion (subjectively interpreted arousal) that is derived from arousal cannot be indicated from any EDA parameter. Therefore, also verbal scales were used. Both, the total amplitude and the frequency show significant correlation with the emotion “joy” (Table 2).

Furthermore, in the experimental store (85.5%), significantly ($p=0.000$), more fruits and vegetables were bought than in control store (46.6%).

Study 2 (replication study for non-food products) was conducted in a shopping mall, where we set up a bookstand in

Table 1
Arousal reactions: experiment 1

Items	Store	Mean (for illustration)	<i>N</i>	Mean rank	Mann–Whitney	Monte Carlo significance
Total amplitude	Experimental	12.753 μ S	15	17.07	44.000	.024
	Control	6.752 μ S	12	10.17		
Frequency	Experimental	84.73	15	17.20	42.000	.016
	Control	47.25	12	10.00		

front of the shop window of a bookstore. Different versions of the bookstand were designed. The first version (C) simply has a long table with many books piled on it, on topics about foreign countries, vacation trips and books for vacation reading. In the second group, the table was decorated with sand and shells, a straw hat, beach towels, and a vivid-looking, attractive mannequin in a deckchair (affective stimuli A). The third variant (A and I) provided the same stimuli as the second group, but here, also posters signaling a bargain in flashy colors were added, in an attempt to attract consumers with extremely low prices. The bargain posters belong to the category of intense stimuli. A “marker” was set to the EDR curve when the test subject reached the bookstand and another one when he or she left. Later on, we analyzed only that part of the EDR curve, which had been recorded in between the two markers. Between the groups, there were no significant differences in the socio-demographic data.

The results show that both parameters, total amplitude and frequency, are highest in the group in which test persons were exposed to all stimuli (A and I), whereas the parameters are lowest in the control group (C), in which only books were presented. Behavioral reactions were tested by a “coupon-question”. Here, test persons were asked to imagine that they received a coupon worth 100€, which they could cash in any bookstore they liked. The subjects were then asked to name the share of the coupon value they would be willing to spend in the specific book store, where we conducted the study. Again, we found significant differences ($p < 0.005$) between the control group on the hand ($M_C = 64\text{€}$), the group with affective stimuli ($M_A = 79\text{€}$), and the special-offer group ($M_{AI} = 88\text{€}$). The design of the bookstand was evaluated best in the 2nd group (A); price image was best in the 3rd group (A and I).

7.2. Studies 3 and 4

Studies 3 and 4 were conducted in 2002, Study 3 on the first floor of a local, medium-sized German shopping mall. Study 4 took place in a large department store for sports goods. For both groups, randomly chosen customers entering the mall or the store were asked to participate in our study. Respondents of all age, income, and education classes were involved. The electrodes were attached to one hand to register the EDR. Both the telemetric and registering devices were conveniently packed in a shoulder bag so that the respondents could walk through the retail environment unhindered by the technical equipment (see Fig. 2).

For a start, participants were asked some “ice-breaking questions”. In the meantime, they got used to the electrodes and the individual baseline of the EDR curve reached “zero” so that subsequently, all phasic arousal reactions could be measured from the same starting-point, thus, rendering EDR curves of different participants comparable. Next, the test persons were instructed to walk up and down the mall or the department store as if they were on a normal shopping trip. They were asked to raise their right arm to signal when they no



Fig. 2. Test person in the mall experiment.

longer felt like shopping or when they had decided to choose a product. Then, electrodes were immediately detached, and test persons were asked to report their product choice. In both studies, time markers were set when the participants started and ended the shopping trip, so as to detect the exact EDR-registration time. After the exclusion of artifacts (e.g. pressure on the electrodes, for details, see Boucsein [5]), we obtained 82 valid EDA data sets for Study 3 (mall) and 105 for Study 4 (department store).

In both studies, respondents were asked whether or not they decided to buy something in order to identify “buyers” and “non-buyers”. The products they decided to buy were noted on the questionnaire. In Study 4, we also controlled whether the selected product had really been bought subsequent to the experiment. This is an important difference between the two experimental settings. In the mall-study, we were only able to establish test persons’ intention to buy something, since, on the first-floor of the mall with its more than 25 different shops, 4 main entrances, and crowded floors, etc., it would have been impossible to observe test persons’ shopping behavior after the interview. In contrast to the situation in the mall, in the third study, the cashier zone, and thus, actual purchases of participants could be observed easily and without problems. Consequently, only in our second study, “real” buying decisions were registered. Furthermore, in the third study, some participants mentioned products that could not be bought in shops on the first floor of the mall (like computers or washing machines). We assumed these

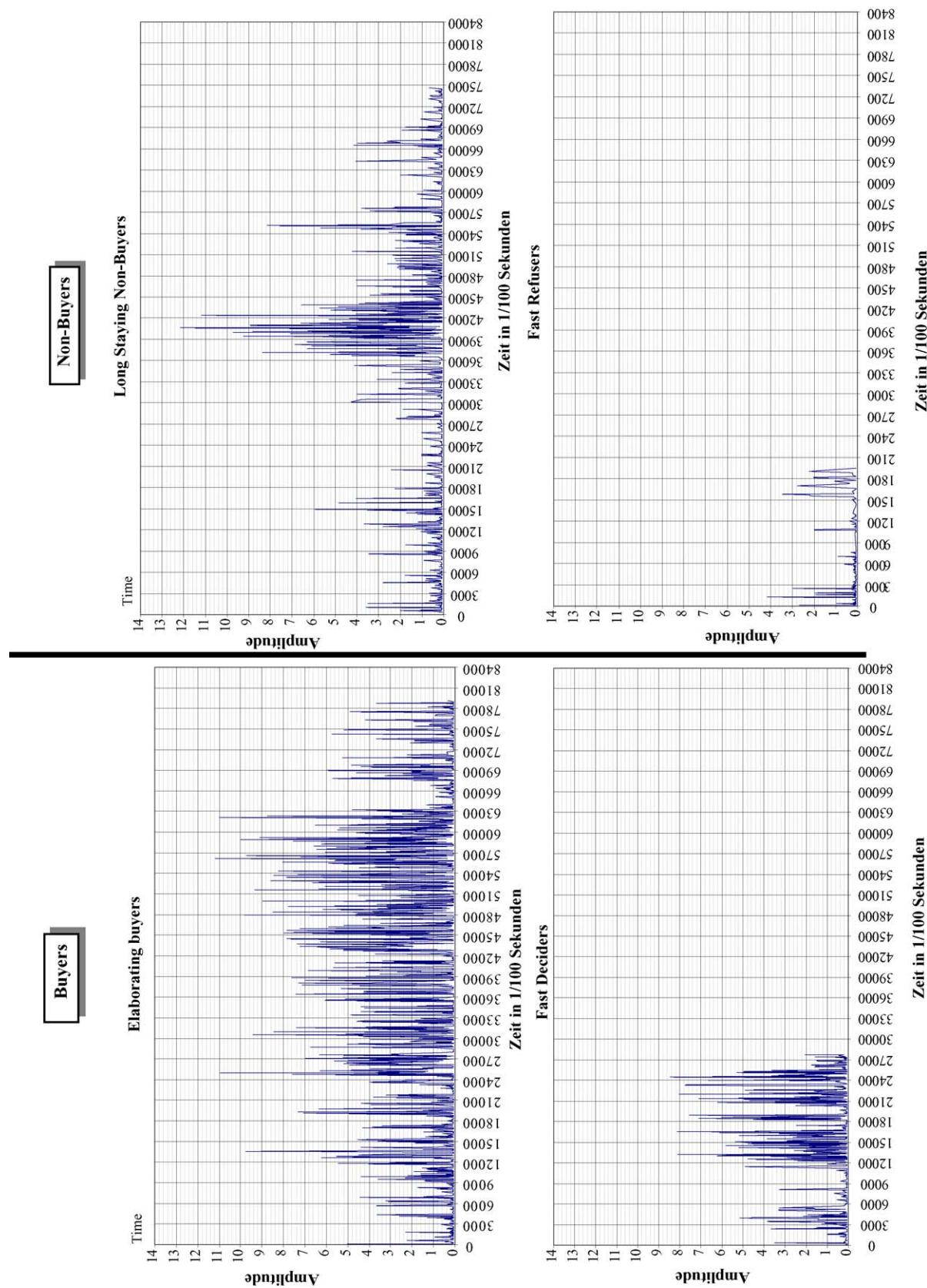


Fig. 3. Typical arousal curves.

Table 3
Phasic arousal of “buyers” and “non-buyers”

EDA parameters	Group	Mean	Mean rank	n	Non-parametric or parametric test	Significance
Study 3						
Total amplitude	Non-buyers (mall)	698.59 μ S	27.52	44	Chi-square; n < 30	Kruskal-Wallis
	Buyers (mall)	1036.73 μ S	38.69	16	4.795	.029
Frequency						
	Non-buyers (mall)	971.64	28.26	44	2.711	n.s.
	Buyers (mall)	1237.50	36.66	16		
Study 4						
Total amplitude	Non-buyers (department store)	614.75 μ S	–	70	F-value (n > 30)	ANOVA
	Buyers (department store)	1136.39 μ S	–	35	9.522	.003
Frequency						
	Non-buyers (department store)	1040.21	–	70	18.938	.000
	Buyers (department store)	1751.31	–	35		

persons had already made their decision at home. Therefore, these cases were excluded from further analysis. Most “buyers” in the first study tended to buy convenience goods (like chocolate, cake, flowers, etc.). As shown in Table 3, in both studies, a comparison between buyers and non-buyers reveals significant differences with respect to total amplitude, thus, supporting the hypothesis. In the second study, the frequency of EDR reactions also varied significantly (for details, see Ref. [25]; Table 3).

In the 3rd study, there was no significant difference between buyers and non-buyers with regard to time spent in the mall (on average, EDR-registration lasted about 6 min), whereas in the department store, buyers (14 min) spent 6 min longer than non-buyers (8 min). In both settings, the experimental design gave test persons the opportunity to expose themselves to any stimuli as long as they liked. This means the individual may gather more EDR responses, the longer

he/she stays at the POS, irrespective of the decision to buy something (or not to). According to the theory, time spent at the POS is dependent on arousal evoked by the store atmosphere [11,38].

To investigate different arousal patterns in the sports department store, also data mining–cluster analyses were conducted with total amplitude, frequency and time as cluster variables. Within these segments, different arousal patterns can be detected (Fig. 3). In a second step, statements measuring buying behavior, where compared between the four groups (see Fig. 4). The results can explain possible causes of the different arousal pattern. “Elaborating buyers” declared that before they decided to buy something they examined if the products matched their life-styles. “Long-staying non-buyers” forgot the time while browsing (therefore, they could also be characterized as “browsers”) and “fast deciders” liked the special-offers and cheap prices. In contrast, “fast refusers”

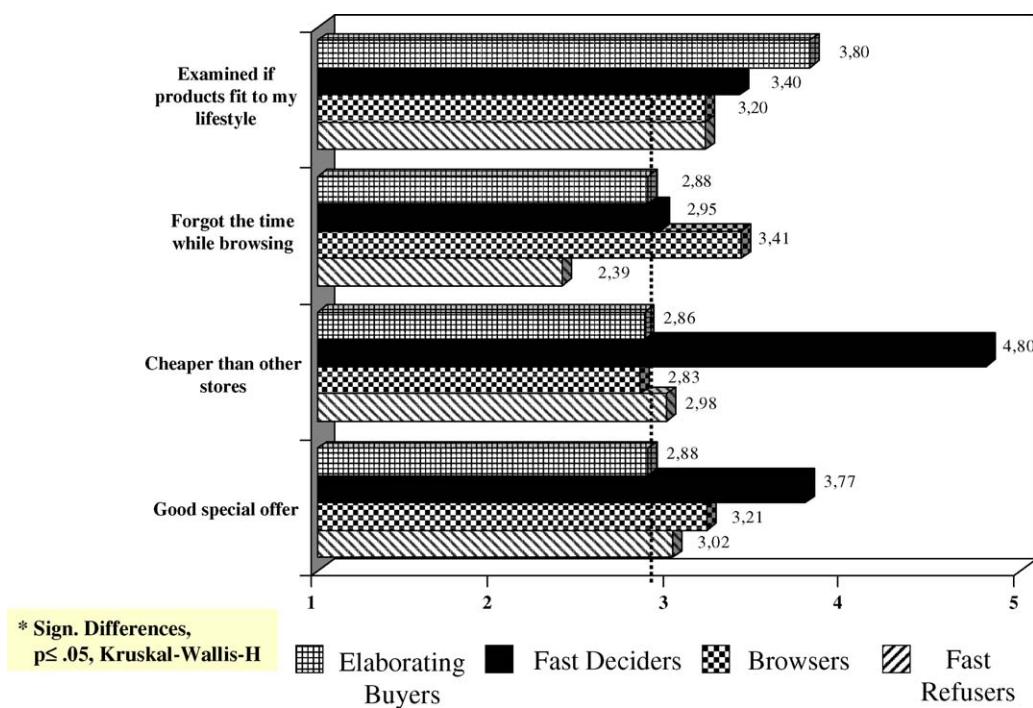


Fig. 4. Attitudes of buyers and non-buyers.

can be characterized by a more negative attitude towards the store.

8. Discussion

The results show that recording EDA at the point-of-sale is a practicable way to measure arousal in a valid manner. The findings of Study 1 confirmed the hypothesis that the experimental store represents the higher arousing store environment than the control store and that arousal was significantly correlated with "joy" (similar results in Study 2). EDR thus seems to be a future research method to work out visual merchandising concepts that evoke an optimum level of arousal. EDR recording can also be applied to explore arousal and orienting reflexes derived from different environmental stimuli settings, and the arousal potential of elements designed to be visually striking. Furthermore, EDR records arousal evoked at the point-of-sale simultaneously to the perception of stimuli during shopping.

Summarizing the essence of Studies 2 und 3, it is evident that arousal is an important construct for the explanation of buying behavior. EDR shows that buyers are more aroused than non-buyers. Moreover, clusters with different arousal patterns can be detected that also show different decision-making patterns. Thus, arousal measurement will be useful to distinguish different problem-solving types (see Ref. [25]).

Furthermore, it would be fascinating to combine marketing relevant results from arousal research at the point-of-sale with fMRI or PET research in the lab [10]. An interesting research question could be to find out first which products evoke high phasic arousal reactions at the point-of sale and then to analyze via brain-imaging, which areas in the brain are activated by these products.

However, EDA recording has also limitations, especially, concerning the time that is needed in order to collect satisfying sample sizes. Also, as long as computerized artifact detection is not possible, artifact detection needs to be done manually by screening each single curve. Another restriction is the fact that EDA cannot reveal whether an arousing situation was perceived with positive or with negative emotions or attitudes, whether a situation was rather pleasing or unpleasing. Verbal control of perceived emotions and attitudes by a standardized questionnaire, therefore, is still necessary.

References

- [1] B.J. Babin, W.R. Darden, M. Griffin, Work and/or fun: measuring hedonic and utilitarian shopping value, *J. Consum. Res.* 10 (1994) 644–656.
- [2] R.P. Bagozzi, The role of psychophysiology in consumer research, in: T.B. Robertson, H.H. Kassarjian (Eds.), *Handbook of Consumer Behavior*, Prentice Hall, Englewood Cliffs, 1991, pp. 124–161.
- [3] R.P. Bagozzi, M. Gopinath, P.U. Nyer, The role of emotions in marketing, *J. Acad. Market. Sci.* 27 (1999) 184–206.
- [4] S. Bekmeier, *Nonverbale Kommunikation in der Fernsehwerbung*, Physica, Heidelberg, 1989.
- [5] W. Boucsein, *Electrodermal Activity*, Plenum Press, NY/London, 1992.
- [6] W. Boucsein, Aktivierung, in: H. Luczak, W. Volpert (Eds.), *Handbuch Arbeitswissenschaft*, Schaeffer-Poeschel, Stuttgart, 1997, pp. 309–312.
- [7] W. Boucsein, Physiologische Grundlagen und Meßmethoden der dermalen Aktivität, in: F. Rösler (Ed.), *Enzyklopädie der Psychologie, Bereich Psychophysiologie*, Band 1: *Grundlagen und Methoden der Psychophysiologie*, Kapitel, Vol. 7, Hogrefe, Göttingen, 2001, pp. 551–623.
- [8] R. Buck, *Human Motivation and Emotion*, second ed., John Wiley and Sons, NY, 1988.
- [9] J.F. Cohn, A.J. Zlochower, J. Lien, T. Kanade, Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding, *Psychophysiology* 36 (1999) 35–43.
- [10] H.D. Critchley, R. Elliott, C.J. Mathias, R.J. Dolan, Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study, *J. Neurosci.* 15, 20 (8) (2000) 3033–3040.
- [11] R.J. Donovan, J.R. Rossiter, Store atmosphere: an environmental psychology approach, *J. Retail.* 28 (1982) 34–57.
- [12] R.J. Donovan, J.R. Rossiter, G. Marcolyn, A. Nesdale, Store atmosphere and purchasing behavior, *J. Retail.* 70 (3) (1994) 283–294.
- [13] E. Duffy, *Activation and Behavior*, John Wiley and Sons, NY/London, 1962.
- [14] P. Ekman, W.V. Friesen, *Facial Action Coding System (FACS): A Technique for the Measurement of Facial Actions*, Consulting Psychologists Press, Palo Alto, 1978.
- [15] P. Ekman, W.V. Friesen, S.S. Tomkins, Facial affect scoring technique: a first validity study, *Semiotica* 3 (1) (1971) 37–58.
- [16] M.W. Eysenck, *Attention and Arousal*, Springer, Berlin, 1982.
- [17] E. Fischer, S.J. Arnold, More than a labor of love: gender roles and Christmas shopping, *J. Consum. Res.* 17 (1990) 333–345.
- [18] M.H. Flicker, W.C. Speer, Emotional responses to store layout and design: an experimental approach, in: A. Parasuraman (Ed.), *Enhancing Knowledge Development in Marketing*, Vol. 1, AMA Educators' Proceedings, AMA, Chicago, 1990, pp. 1–5.
- [19] M. Fredrikson, T. Furmark, M.T. Olsson, H. Fischer, J. Andersson, B. Langström, Functional neuroanatomical correlates of electrodermal activity: a positron emission tomographic study, *Psychophysiology* 35 (1998) 179–185.
- [20] A. Groeppel-Klein, The influence of the dominance perceived at the PoS on the price-assessment, in: B.G. Englis, A. Olofsson (Eds.), *European Advances in Consumer Research*, Vol. 3, Association for Consumer Research, Provo, UT, 1997, pp. 304–311.
- [21] A. Groeppel-Klein, The avoidance of usury-images evoked by exclusive store-designs, in: E. Breivik, A.W. Falkenberg, K. Grønhaug (Eds.), *Rethinking European Marketing, Proceedings From the 30th EMAC conference*, The Norwegian School of Economics and Business Administration, Bergen, 2001, pp. 193–200.
- [22] A. Groeppel-Klein, Aktivierungsforschung und Konsumentenverhalten, in: A. Groeppel-Klein (Ed.), *Konsumentenverhaltensforschung im*, Vol. 21, Jahrhundert, Wiesbaden, 2004, pp. 29–66.
- [23] A. Groeppel-Klein, D. Baun, The role of customers' arousal for retail stores—results from an experimental pilot study using electrodermal activity as indicator, in: M.C. Gilly, J. Meyers-Levy (Eds.), *Adv. Consum. Res.* XXVIII (2001) 412–419.
- [24] A. Groeppel, B. Bloch, An investigation of experience-orientated consumers in retailing: the international review of retail, Distrib. Consum. Res. 1 (4) (1990) 101–118.
- [25] A. Groeppel-Klein, C.C. Germelmann, A. Domke, H. Woratschek, Arousal as a driving force for decision making—empirical results from measuring electrodermal reactions at the point-of-sale, *Adv. Consum. Res.* XXXII, in press.
- [26] M.B. Holbrook, E.C. Hirschmann, The experiential aspects of consumption: consumer fantasies, feelings, and fun, *J. Consum. Res.* 9 (1982) 132–140.

- [27] W. Kroeber-Riel, Activation Research—psychobiological approaches in consumer research, *J. Consum. Res.* 5 (1979) 240–250.
- [28] J. Le Doux, *The Emotional Brain. The Mysterious Underpinnings of Emotional Life*, Simon, Schuster, NY, 1996.
- [29] D.B. Lindsley, Emotions, in: S.S. Stevens (Ed.), *Handbook of Experimental Psychology*, Wiley, NY, 1951, pp. 473–516.
- [30] A. Mehrabian, J.A. Russell, *An Approach to Environmental Psychology*, MIT, Cambridge, MA, 1974.
- [31] G. Meyer-Hentschel, *Aktivierungswirkung von Anzeigen*, Physica, Würzburg, Wien, 1983.
- [32] C.M. Portas, G. Rees, A.M. Howseman, O. Josephs, R. Turner, C.D. Frith, A specific role for the thalamus in mediating the interaction of attention and arousal in humans, *J. Neurosci.* 18 (21) (1998) 8979–8989.
- [33] J.A. Russell, A. Mehrabian, Evidence for a three-factor theory of emotions, *J. Res. Perspect.* 11 (1977) 273–294.
- [34] J.F. Sherry Jr., A sociocultural analysis of a midwestern flea market, *J. Consum. Res.* 17 (1990) 13–30.
- [35] A. Steiger, *Computergestützte Aktivierungsmessung in Der Marketingforschung*, Peter Lang, Frankfurt am Main, Bern, NY/Paris, 1988.
- [36] S.H.C. Tai, A.M.C. Fung, Application of an environmental psychology model to in-store buying behaviour: *The International Review of Retail, Distrib. Consum. Res.* 7 (4) (1997) 311–337.
- [37] Y. Tian, T. Kanade, J.F. Cohn, Recognizing action units for facial expression analysis, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2) (2001) 97–115.
- [38] P. van Kenhove, P. Desrumaux, The relationship between emotional states and approach or avoidance responses in a retail environment: *The International Review of Retail, Distrib. Consum. Res.* 7 (1997) 351–368.
- [39] P.H. Venables, M.J. Christie, Eectrodermal activity, in: I. Martin, P.H. Venables (Eds.), *Techniques in Psychophysiology*, Wiley, NY, 1980, pp. 3–67.
- [40] P. Vitouch, Psychophysiological methods in media research, in: P. Winterhoff-Spurk, T.H.A. van der Voort (Eds.), *New Horizons in Media Psychology*, Westdeutscher Verlag, Opladen, 1997, pp. 116–124.

Neural Correlates of Behavioral Preference for Culturally Familiar Drinks

Samuel M. McClure,^{1,2} Jian Li,¹ Damon Tomlin,
Kim S. Cypert, Latané M. Montague,
and P. Read Montague*

Department of Neuroscience
Menninger Department of Psychiatry
and Behavioral Sciences
Baylor College of Medicine
1 Baylor Plaza
Houston, Texas 77030

Summary

Coca-Cola® (Coke®) and Pepsi® are nearly identical in chemical composition, yet humans routinely display strong subjective preferences for one or the other. This simple observation raises the important question of how cultural messages combine with content to shape our perceptions; even to the point of modifying behavioral preferences for a primary reward like a sugared drink. We delivered Coke and Pepsi to human subjects in behavioral taste tests and also in passive experiments carried out during functional magnetic resonance imaging (fMRI). Two conditions were examined: (1) anonymous delivery of Coke and Pepsi and (2) brand-cued delivery of Coke and Pepsi. For the anonymous task, we report a consistent neural response in the ventromedial prefrontal cortex that correlated with subjects' behavioral preferences for these beverages. In the brand-cued experiment, brand knowledge for one of the drinks had a dramatic influence on expressed behavioral preferences and on the measured brain responses.

Introduction

Perceptual constructs are generally multidimensional, integrating multiple physical and cognitive dimensions to generate coherent behavioral preferences. In sensory processing, the idea of multidimensional integration has long been used to frame a range of questions about cross-modal interactions in physiological and behavioral responses (Stein et al., 1996; 1999; Wallace and Stein, 1997; Armony and Dolan, 2001; Dolan et al., 2001; Laurienti et al., 2002, 2003). This same multidimensional perspective has also been developed for olfactory and gustatory processing, where the detection, discrimination, and perceived intensity of stimuli are not only functions of the primary physical properties (odors, flavors) but are also modulated "cross-modally" by visual input (Gottfried and Dolan, 2003), auditory input, and current reward value (Gottfried et al., 2003).

The work just described has focused on the perceptual discrimination of odors and flavors, the correlated

neural responses, and the modulation of both by nonodor or nonflavor stimuli—that is, the sensory problem. Ultimately, such sensory discriminations and the variables that influence them serve to influence expressed behavioral preferences. Hence, there is another large piece of the problem to understand. For modern humans, behavioral preferences for food and beverages are potentially modulated by an enormous number of sensory variables, hedonic states, expectations, semantic priming, and social context. This assertion can be illustrated with a quote from Anderson and Sobel (2003) profiling the work of Small et al. (2003) on taste intensity and pleasantness processing:

"A salad of perfectly grilled woodsy-flavored calamari paired with subtly bitter pale green leaves of curly endive and succulent petals of tomato flesh in a deep, rich balsamic dressing. Delicate slices of pan-roasted duck breast saturated with an assertive, tart-sweet tamarind-infused marinade."

The text goes on further, but note that the sheer lushness of the description adds somehow to the appeal of the food described. Also notice one implicit point of the description: many levels of social, cognitive, and cultural influences combine to produce behavioral preferences for food and drink. The above description likely would not appeal to a strict vegan or an owner of a pet duck. Anderson and Sobel point out that the preferences indexed by their prose originated from the economic demands on our early forebears and were unlikely to have been strictly about aesthetic responses to food and drink.

However, the modern problem is different. Cultural influences on our behavioral preferences for food and drink are now intertwined with the biological expediency that shaped the early version of the underlying preference mechanisms. In many cases, cultural influences dominate what we eat and drink. Behavioral evidence suggests that cultural messages can insinuate themselves into the decision-making processes that yield preferences for one consumable or another. Consequently, the appeal or repulsion of culturally relevant sights, sounds, and their associated memories all contribute to the modern construction of food and drink preferences. The neural substrates underlying food and drink preferences and their influence by cultural images have not been explored. As alluded to above, the majority of work on olfaction and gustation has focused on sensory processing. In this paper, we combine simple taste tests and event-related functional magnetic resonance imaging (fMRI) to probe the neural responses that correlate with the behavioral preference for noncarbonated versions of Coke® and Pepsi®. We further investigate the influence of the brand image on behavioral choice and brain response to both drinks.

These two stimuli were chosen for three reasons. (1) They are culturally familiar to subjects. (2) They are both primarily composed of brown, carbonated sugar water, and sugar water serves as a primary reward in many

*Correspondence: read@bcm.tmc.edu

[†]These authors contributed equally to this work.

²Present address: Department of Psychology, Princeton University, Princeton, New Jersey 08544.

animal and human experiments. (3) Despite their similarities, they generate a large subjective preference difference across human subjects, which might correlate with fMRI-measured brain responses. We pursued three primary questions using the experiments presented in this paper. (1) What is the behavioral and neural response to these drinks when presented anonymously? (2) What is the behavioral and neural influence of knowledge about which drink is being consumed? (3) In questions 1 and 2, is there a correlation between the expressed behavioral preference and the neural response as measured using fMRI?

The medical importance of understanding these questions is straightforward—there is literally a growing crisis in obesity, type II diabetes, and all their sequelae that result directly from or are exacerbated by overconsumption of calories (for recent work see Chacko et al., 2003; Ford et al., 2003; Wyatt, 2003; Zimmet, 2003; Popkin and Nielsen, 2003). It is now strongly suspected that one major culprit is sugared colas (Popkin and Nielsen, 2003). The possibility of obtaining coherent answers to these questions derives from the growing fMRI work on reward processing.

Recent work using fMRI has identified reward-related brain responses that scale with the degree to which subjects find stimuli pleasing or rewarding (Knutson et al., 2001; Aharon et al., 2001). With this information in hand, it is tempting to suggest that humans will choose more pleasing stimuli over less pleasing stimuli by evaluation and comparison and that, for our two sugared drinks, the most pleasing drink is the one that subjectively tastes better than its competitor. This perspective offers the simplest model that connects reward-related brain responses to expressed behavioral preferences. However, most real-world settings present numerous primary sensations and top-down influences that act to organize a coherent behavioral preference. Studies have indeed shown that cultural information can modulate reward-related brain response (Erk et al., 2002). This general observation is particularly true for Coke and Pepsi; that is, there are visual images and marketing messages that have insinuated themselves into the nervous systems of humans that consume the drinks. It is possible that these cultural messages perturb taste perception; however, no direct neural probes of this possibility have been carried out. It is this issue and its implications that we sought to address, and our results suggest that there might be parallel mechanisms in the brain cooperating to bias preference.

Results

A total of 67 subjects participated in the study. They were separated into four groups ($n_1 = 16$, $n_2 = 17$, $n_3 = 16$, $n_4 = 18$). Each group was given a separate taste test outside the scanner and drink delivery paradigm while in the scanner.

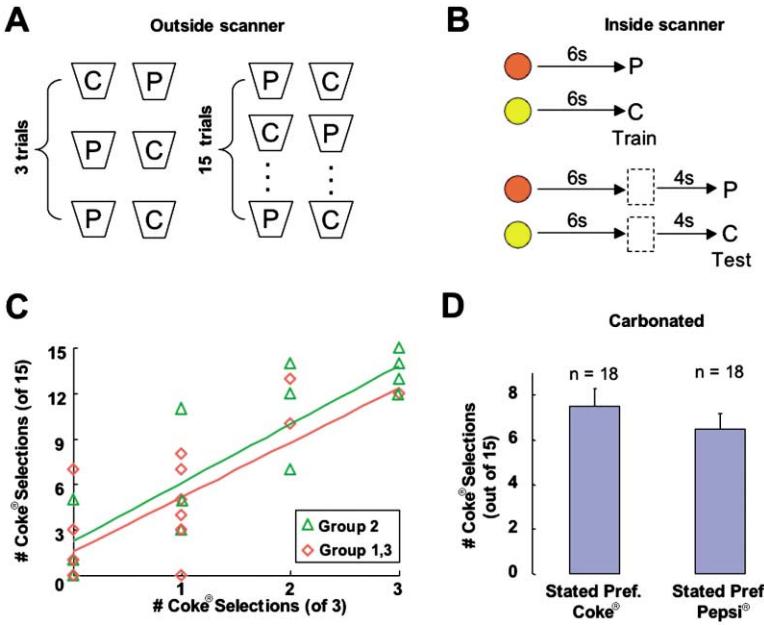
For all four groups of subjects, two separate measures of behavioral preferences were obtained. First, subjects were asked “Which drink you prefer to consume: Coke, Pepsi, or no preference?” Their answers are referred to as their *stated preferences*. Next, subjects engaged in three rounds of a forced-choice taste test (see Experi-

mental Procedures for details). In the taste test, the first two groups chose between two unmarked cups, one of which contained Pepsi and the other Coke in a 3-trial and a 15-trial behavioral task, correspondingly (*anonymous taste test*; Figure 1A). The other two groups (*semianonymous taste tests*) made three preference decisions, but in this case both cups contained the same drink (either Pepsi or Coke); however, one cup was unlabeled and the other indicated the brand of the drink contained in the cup. In these semianonymous taste tests, subjects were told that the unlabeled cups contained either Pepsi or Coke, and hence no deception was involved. Preferences exhibited during the taste tests are referred to as *behavioral preferences*.

To determine how subjects’ preferences interact with brand information at the level of brain activity, the four subject populations completed scanning experiments analogous to the taste test manipulations. The taste test outside the scanner and drink delivery paradigm inside the scanner are illustrated in Figures 1A and 1B. The design of these experiments was complicated by two facts: (1) blood oxygenation level-dependent (BOLD) responses in reward-related brain areas are significantly affected by whether stimuli are predictable in time (Berns et al., 2001; O’Doherty et al., 2003a; McClure et al., 2003), and (2) we wished to study differential brain responses to soda delivery alone. Because of these constraints, we could not show brand information simultaneous with soda delivery since this would make it impossible to separate the perception of brand information from the response to soda delivery. However, with brand information displayed prior to soda delivery, we had to contend with making the soda delivery predictable. There were two possible strategies: (1) randomize the time between the display of brand information and delivery of soda or (2) construct a design in which brain responses can still be elicited when brain information consistently (fixed time) precedes soda delivery. Since the effect of unpredictability is still a matter of investigation and there is some evidence that brain responses may be elicited in reward-related brain structures based on unpredictability alone (Zink et al., 2004), we opted for the second option. In particular, we adopted the design of McClure et al. (2003) in which strong BOLD responses were elicited even after a large number of event repetitions.

We trained subjects to expect Coke and Pepsi at fixed times (6 s) following distinct visual cues (Figures 1B, 4A, and 4C, *training period*). After training, we then studied brain responses evoked by delayed, unexpected (10 s following cue) cola delivery. For the first and second subject groups (*anonymous*), the predictive visual cues were flashes of yellow and red light, counterbalanced and paired with subsequent Coke and Pepsi delivery. In the third and fourth subject groups (*semianonymous*), the two fluids were the same (both either Pepsi or Coke). One of the cues was anonymous (yellow or red light), and the other contained brand information (picture of a Coke can or a Pepsi can; Figures 4A and 4C).

In all fMRI experiments in which subjects must swallow, head motion is a concern. Using head constraints, subjects’ head movements remained no larger than 2 mm during the entire experiment. In addition, we performed two separate analyses that both ruled out the



(D) Average Coke preference in Coke and Pepsi drinkers in a 15-trial carbonated Coke-Pepsi taste task. Average Coke selection is 7.5 ± 0.8 (mean \pm SE) for Coke drinkers and 6.8 ± 0.5 (mean \pm SE) for Pepsi drinkers. Our data do not support that stated preference is correlated with behavioral preference in the carbonated state (two-tailed Student's t test, $p = 0.46$, $n_1 = 18$, $n_2 = 18$).

possibility that any of the following results were contaminated by this possible confound (see Experimental Procedures for more details). First, a two-way ANOVA test of the event-related head movements and six parameters of head movements (x , y , z , pitch, yaw, roll) in all of the four groups was performed. We found that movement was uncorrelated with any event ($p > 0.6$ for all movement parameters). Second, in those areas which are identified as "active," we used head movement parameters as regressors to determine whether their activities were significantly correlated with head movements. This revealed no statistical significance ($p > 0.1$ for all movement parameters).

All of the subsequent findings relate to brain structures that are not directly involved in processing gustatory stimuli. Therefore, it is important to note that significant brain activity was evoked by the delivery of Coke or Pepsi in gustatory cortical regions (insular cortex; $p < 0.01$ for both drinks) but was unaffected by any of our experimental factors.

Group 1-2: Anonymous Taste Test

The results of the anonymous taste test (3-trial version, group 1) are shown in Figure 3A. A histogram is shown which summarizes the behavioral preferences for all 16 subjects. Subjects were balanced, with a nearly equal number preferring Pepsi, preferring Coke, or showing no distinct preference. Similarly, there was no difference in subjects' stated preferences (data not shown), with an equal number of subjects declaring a preference for Coke ($n = 7$) and Pepsi ($n = 6$; $p = 0.79$). However, the correlation between subjects' stated and behavioral preferences does not reach statistical significance ($r^2 = 0.14$; $p = 0.16$).

Figure 1. Anonymous Coke and Pepsi Task Settings and Behavioral Results

(A) Anonymous taste test. Each subject was given a taste test outside the scanner. The test required subjects to make 3 separate choices (groups 1 and 3) or 15 choices (group 2) in which they indicated their preference for the soda in one of two unmarked cups. One of the cups contained 10 mL of Coke, and the other contained an equal volume of Pepsi. (B) In the scanner, subjects were trained to expect soda delivery at a 6 s delay following light illumination using a Pavlovian conditioning paradigm. Twenty light-drink pairings were used for training, separated by random time intervals of between 6 and 16 s. Following training, 6 test pairings for each liquid were randomly interleaved during the succeeding 25 pairings in which soda delivery was unexpectedly delayed for 4 s; evoked brain responses were studied for these 12 delayed deliveries.

(C) Correlation of Coke preference in behavioral taste tasks between the original 3 trials and later 15 trials (groups 1 and 3; red; $r^2 = 0.51$, $n = 15$), between first 3 trials and the whole 15 trials in the other independent 15-trial anonymous taste test (group 2; green; $r^2 = 0.78$, $n = 16$).

For technical reasons (carbonation builds up in delivery tubes causing unreliable soda delivery in the scanner experiment), all the taste tests were originally conducted with decarbonated soft drinks. However, we now show that the behavioral results are unaffected by carbonation. A separate anonymous taste task of 15 forced-choice trials was presented to each subject with carbonated drinks (Coke and Pepsi). Stated and behavioral preferences were not correlated in this condition (Figure 1D, two-tailed Student's t test, $p = 0.46$).

Behavioral preferences measured in the 3-trial taste task are a potentially unreliable measure of subjects' true preferences due to the small number of measurements involved in the test. To account for this, we recalled our subjects at a delay of several months and had them repeat the taste test with 15 trials. The outcome of these two separate tests are strongly correlated, as shown in Figure 1C (subjects recalled from groups 1 and 3; red; $y = 3.6x + 1.5$, $r^2 = 0.51$, $n = 15$). Consistency of preferences in 3-trial and 15-trial taste tests was further confirmed within session. In a separate group of subjects (Figure 1C, green), the outcome of their first 3 trials well predicted the outcome of the full 15-trial test ($y = 3.8x + 2.2$, $r^2 = 0.78$, $n = 16$).

Group 1-2: Scanning, Anonymous Drink Delivery

A linear regression analysis using behavioral preferences from the 3-trial anonymous taste task as a regressor indicated that the difference in brain responses evoked by Coke and Pepsi in the ventromedial prefrontal cortex (VMPFC; MNI coordinates [8, 56, 0]; peak z score 3.44) (Figure 2B) scaled monotonically with the results of the behavioral taste test (Figure 2A) (no other significant regions at $p < 0.001$, uncorrected for multiple comparisons).

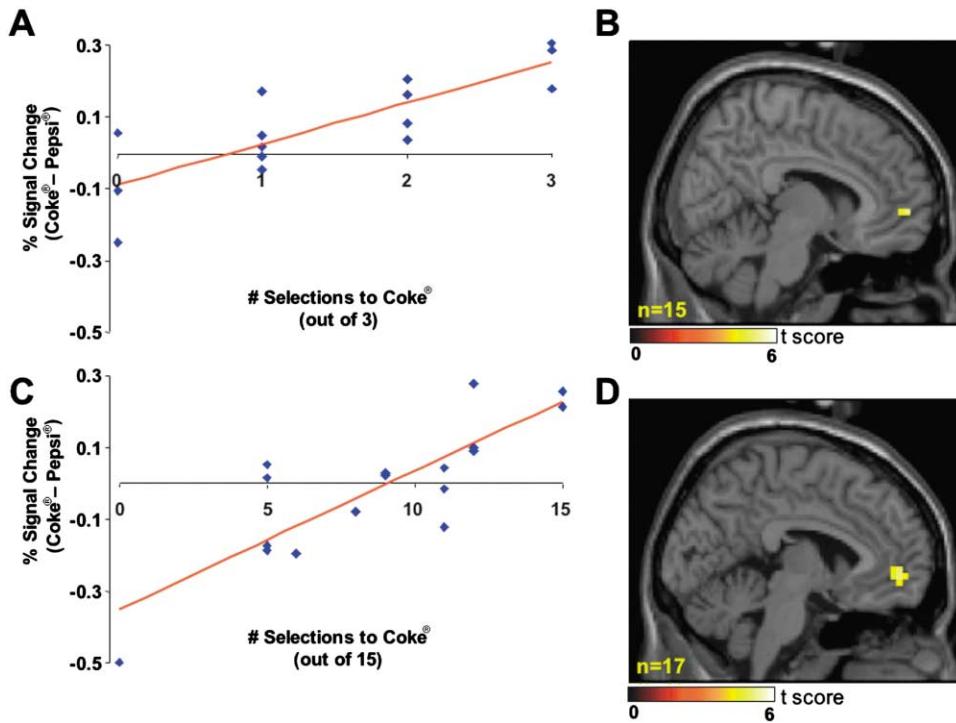


Figure 2. Neural Correlates of Preference for Anonymous Coke and Pepsi Delivery in 3-Trial and 15-Trial Anonymous Taste Tasks

(A) Behavioral preferences expressed in the 3 trial taste test varied linearly with brain responses in the ventromedial prefrontal cortex (group 1). The vertical axis is the contrast (delayed Coke response – delayed Pepsi response) for the voxels shown in (B).
(B) SPM of neural correlates of behavior preference shown in (A) (thresholded at $p < 0.001$; uncorrected for multiple comparisons).
(C) Correlation between behavioral preferences expressed in the 15 trial taste and brain responses in the ventromedial prefrontal cortex (group 2).
(D) SPM of neural correlates of behavior preference shown in (C) (thresholded at $p < 0.001$; uncorrected for multiple comparisons).

sons). As with behavioral preferences, the brain response in the VMPFC was independent of the subjects' stated preferences (paired Student's t test, $p = 0.87$).

As with the behavioral results, it is possible that this finding may suffer from noise in our estimates of subjects' preferences. However, the correlation of BOLD responses in the VMPFC with preference was replicated in the subjects from group 2 whose preference measures were based on 15-round taste tests (Figures 2C and 2D, MNI coordinates [8, 60, 0]; peak z score 3.83; $p < 0.001$, uncorrected). The same region of the VMPFC is also significantly correlated with behavioral preference when data from group 1 and group 2 are combined ($p < 0.001$, uncorrected for multiple comparisons; data not shown).

Individual subjects generally had a strong stated preference for either Coke or Pepsi, and, at any particular time, a guess of which soda they were receiving may have influenced evoked neural responses. The results presented here are not likely to be due exclusively to such top-down influences of brand preference, since stated preference did not correlate with the behavioral preference (taste test results) ($r^2 = 0.14$). However, to explicitly test for effects of brand knowledge, this influence was directly modulated in the following two tasks.

We developed the working hypothesis that the label of either or both drinks would influence the expressed behavioral preference of the subjects. In particular, we tested whether knowledge of which cola was being consumed influenced subjects' responses.

Group 3: Semianonymous Taste Test, Coke

As before, three pairs of cups were presented to the subjects. However, in each pair one of the cups was labeled "Coke" and the other was left unlabeled. For the unlabeled cups, the subjects were told that they could contain either Coke or Pepsi. A Mann-Whitney U test showed that the effect of the Coke label was significant when compared with the anonymous taste test, with subjects showing a strong bias in favor of the labeled cup (Figure 3C, $p < 0.05$). This was not likely to be a result of spurious subject sampling, because, when the subjects were later asked to complete the anonymous taste test, their results were not significantly different from the group 1 (anonymous) results (Figure 3A, $p = 0.84$). Furthermore, these behavioral effects did not correlate with subjects' stated preferences ($p = 0.92$).

Group 3: Scanning, Semianonymous Coke

Figure 4A shows the stimulus paradigm for the Coke label task. In this condition, one cue was a depiction of a Coke can followed 6 s later by Coke delivery. The other stimulus was a light followed by Coke delivery. The number of cue-drink pairings, the number of catch trials per cue, and the pseudorandom times between pairings were exactly the same as in the anonymous drink delivery task described above (Group 1-2: Scanning; see Experimental Procedures). We contrasted the brain response to surprising delivery of Coke when it was known to be Coke with the surprising delivery of

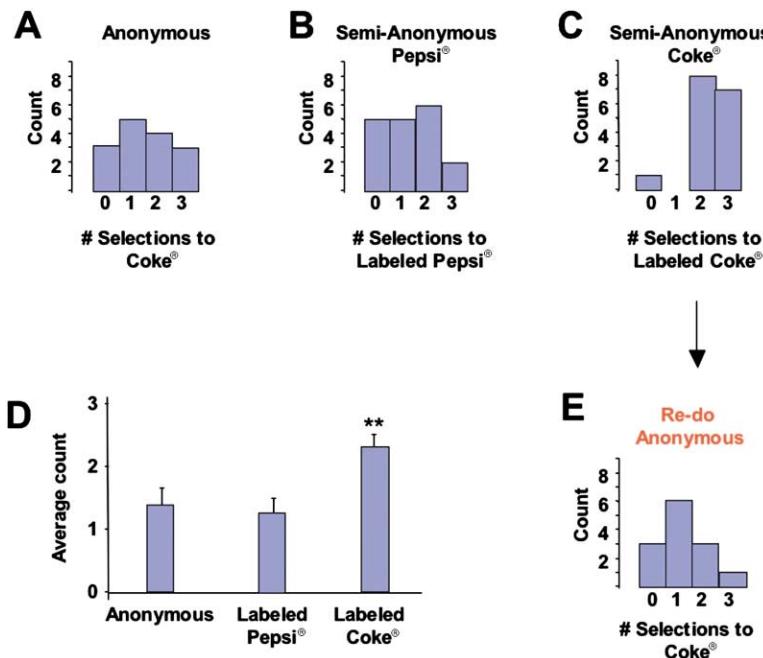


Figure 3. Effect of Brand Knowledge on Behavioral Preferences

(A) Histogram of subjects' preference in double anonymous task. The x axis indicates the number of selections made to Coke (maximum of three). Subjects showed no bias for either Coke or Pepsi.

(B) Histogram of subjects' behavior preference in semianonymous Pepsi task. The x axis indicates the number of selections to the Pepsi-labeled cup. Subjects showed no bias for either the labeled or unlabeled drink.

(C) Histogram of subjects' behavior preference in the semianonymous Coke task. The x axis indicates the number of selections to the labeled Coke. This preference distribution is different from the double anonymous task (Mann-Whitney U task, $n_1 = 16$, $n_2 = 16$, $U = 191.5$, $p < 0.05$) and semianonymous Pepsi task ($n_1 = 18$, $n_2 = 18$, $U = 225.5$, $p < 0.005$), with subjects demonstrating a strong bias in favor of the labeled drink.

(D) Average scores of subjects' preference (number of selections to Coke, labeled Pepsi, and labeled Coke, respectively) in the three behavioral tasks (A-C). Subjects tended to prefer the labeled Coke drink over anonymous Coke (one-way Student's t test, $p < 0.01$).

0.01). The Coke label had a bigger effect in biasing subjects' preferences than the Pepsi label (one-way Student's t test, $p < 0.005$). (E) Subjects who participated in the semianonymous Coke task later completed the anonymous taste test. The distribution of people's preference is significantly different from the Coke-labeled task (Mann-Whitney U test, $n_1 = 16$, $n_2 = 13$, $U = 142.5$, $p < 0.01$) but no different from the results in (A).

Coke when it could have been Coke or Pepsi. The results are shown in Figure 4B. Significant differential activity is observed in several brain areas ($p < 0.001$, uncorrected): bilateral hippocampus, parahippocampus, midbrain, dorsolateral prefrontal cortex (DLPFC), thalamus, and left visual cortex. Details are listed in Table 1. At $p < 0.005$ (uncorrected), the activation in the left hippocampus, left parahippocampus, and midbrain are contiguous (Table 1). BOLD signal changes in the area of the VMPFC identified in the anonymous task were unaffected by brand knowledge (two-tailed paired Student's t test, $p = 0.96$).

Group 4: Semianonymous Taste Test, Pepsi

The taste test for this group was conducted exactly as for group 3 (semianonymous Coke), except both cups in each pair contained Pepsi, and one was labeled as Pepsi. Again, subjects were told that the unlabeled cup could contain either Coke or Pepsi. Unlike the Coke label, the existence of the Pepsi label did not change the distribution of choices significantly relative to the anonymous taste test (Figure 3B, Mann-Whitney U test, $p = 0.82$). Furthermore, selections were biased in favor of the Coke label (in the semianonymous task, above) to a significantly greater degree than they were in favor of the Pepsi label (Figure 3D, $p < 0.005$).

Group 4: Scanning, Semianonymous Pepsi

Figure 4C shows the stimulus paradigm for the Pepsi label task. As with group 3, we contrasted the brain response elicited by the unexpected delivery of labeled versus unlabeled Pepsi. At a threshold of $p < 0.001$ (uncorrected), no brain areas showed a significant main

effect of brand knowledge for Pepsi (Figure 4D). As in the semianonymous Coke task, activity in the VMPFC did not show a significant effect of brand knowledge ($p = 0.89$). Further analysis of the hippocampus and DLPFC revealed that these areas were not significant even at lower thresholds ($p < 0.01$, uncorrected). In particular, the p value within the area of the hippocampus identified in the semianonymous Coke task was 0.43, while in the DLPFC it was 0.41. Further, exclusively masking the results in the semianonymous task (at $p < 0.01$) with the results from the semianonymous Coke task revealed no common areas of activation. Thus, it seems that brand knowledge for Coke and Pepsi have truly different responses both in terms of affecting behavioral preference and in terms of modifying brain responses.

Discussion

In these experiments, we used functional brain scanning to find correlates of people's preferences for two similar sugared drinks: Coke and Pepsi. We report the finding that two separate systems are involved in generating preferences. When judgments are based solely on sensory information, relative activity in the VMPFC predicts people's preferences. However, in the case of Coke and Pepsi, sensory information plays only a part in determining people's behavior. Indeed, brand knowledge (at least in the case of Coke in our study) biases preference decisions and recruits the hippocampus, DLPFC, and midbrain. Our results suggest that the VMPFC and hippocampus/DLPFC/midbrain might function independently to bias preferences based on sensory and cultural information, respectively.

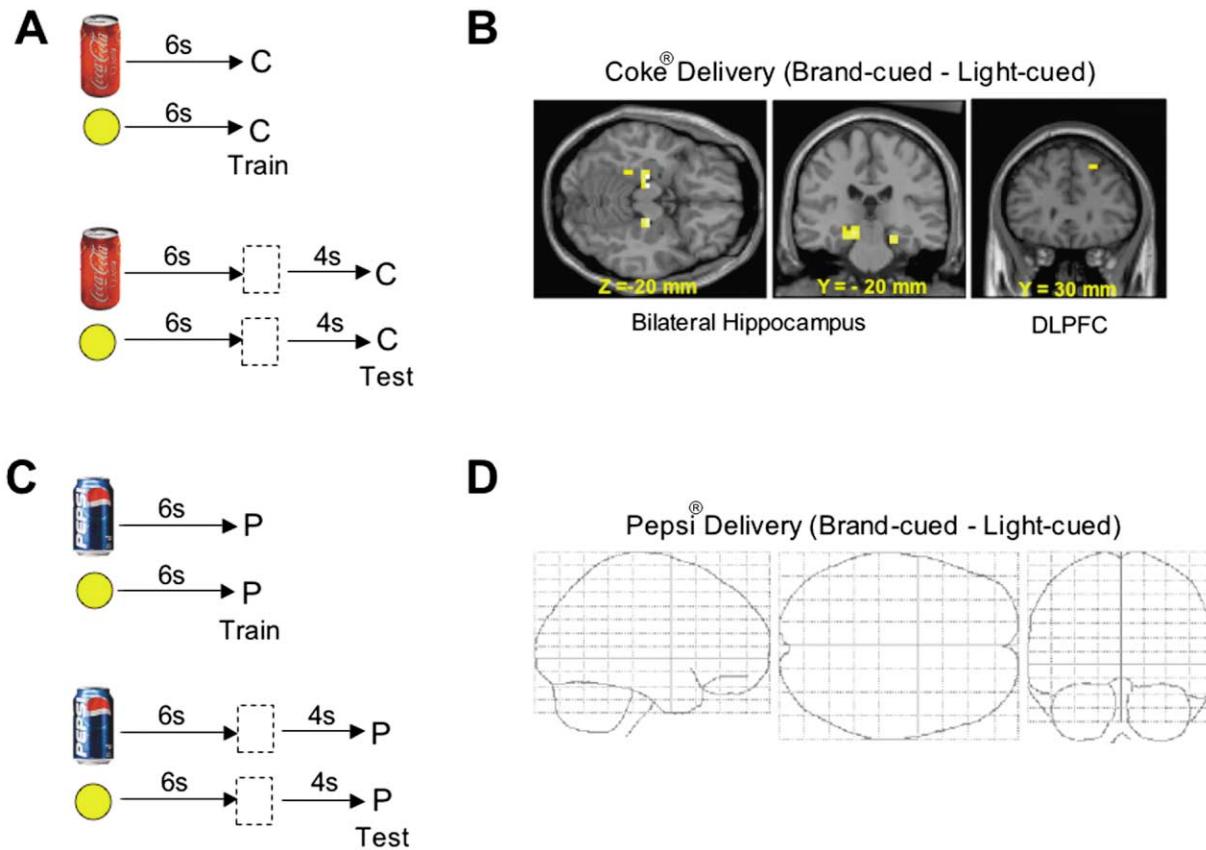


Figure 4. Effect of Brand Knowledge on Brain Responses in Semianonymous Tasks

(A) An image of a Coke can was used to cue the occurrence of Coke. A red or yellow circle (randomized across subjects) predicted the other. Both sodas delivered were Coke.

(B) Coke delivered following an image of a Coke can evoked significantly greater activity in several regions when contrasted against Coke delivered following a neutral flash of light. Significant activations ($p < 0.001$, uncorrected) were found bilaterally in the hippocampus (MNI coordinates $[-24, -24, -20]$ and $[20, -20, -16]$), in the left parahippocampal cortex (MNI coordinates $[-20, -32, -8]$), midbrain (MNI coordinates $[-12, -20, -16]$), and dorsolateral prefrontal cortex (MNI coordinates $[20, 30, 48]$). See Table 1 for details.

(C) In the scanner, an image of a Pepsi can was used to cue the occurrence of Pepsi. A red or yellow circle predicted the other soda, and both sodas delivered were Pepsi.

(D) No voxels survive $p < 0.001$ threshold (uncorrected) for the equivalent contrast in the semianonymous Pepsi experiment.

Coke and Pepsi are special in that, while they have very similar chemical composition, people maintain strong behavioral preferences for one over the other. We initially measured these behavioral preferences objectively, by administering double-blind taste tests. We found that subjects split equally in their preference for Coke and Pepsi in the absence of brand information

(Figure 3A). The functional brain imaging results corroborate the behavioral taste test results. The BOLD signal in the VMPFC correlated strongly with the behavior results of the double-blind taste tests. This area of the brain is strongly implicated in signaling basic appetitive aspects of reward. Imaging data in healthy subjects indicated with BOLD signal changes scale in the VMPFC

Table 1. Location of Brain Areas that Respond Preferentially to Brand-Cued versus Light-Cued Coke Delivery

Brain Region	Peak MNI Coordinates	Peak t Statistic	z Score	Number of Voxels
R hippocampus (BA35)	20, -20, -16	5.37	3.95	7 (30)
L hippocampus (BA35)	-24, -24, -20	5.30	3.91	14
L parahippocampus	-20, -32, -8	4.35	3.45	6 } (72)
Midbrain	-12, -20, -16	6.81	4.53	12
R DLPFC	20, 28, 48	4.79	3.67	3 (8)
R thalamus	12, -8, 4	4.94	3.75	6 (7)
L visual cortex (BA17)	-4, -60, 12	4.38	3.46	4 (19)

Activations are for the semianonymous (Coke) experiment ($p < 0.001$ and $p > 0.005$ in parentheses). L, left hemisphere; R, right hemisphere.

with reward value (Knutson et al., 2001; O'Doherty et al., 2003b). Furthermore, patients with lesions in the VMPFC are insensitive to future reward or punishment value in making decisions (Bechara et al., 1994). A related brain region, the medial orbitofrontal cortex (MOFC), is strongly related to the VMPFC in terms of function. Our imaging parameters resulted in significant signal loss in this area due to magnetic field inhomogeneities. It remains untested, therefore, whether the MOFC shows preference-related responses.

The other special characteristic of Coke and Pepsi is that both possess a wealth of cultural meaning. One important fact that may account for the lack of correlation between subjects' stated and behavioral preferences is the effect of presumed brand knowledge. In accord with this, subjects expressed strong preferences for either Coke or Pepsi when asked which type of soda they normally drink and commonly demonstrated a desire to prove this preference in the anonymous taste test.

To test for effects of brand knowledge, we conducted a series of semianonymous taste tests and imaging experiments. In the taste tests, we found no significant influence of brand knowledge for Pepsi contrasted with the anonymous task. However, there is a dramatic effect of the Coke label on subjects' behavioral preference. Despite the fact that there was Coke in all cups during the taste test, subjects in this part of the experiment preferred Coke in the labeled cups significantly more than Coke in the anonymous task and significantly more than Pepsi in the parallel semianonymous task.

The effects of brand knowledge for Pepsi and Coke were reflected in the imaging experiments as well. When an image of a Coke can preceded Coke delivery, significantly greater brain activity was observed in the DLPFC, hippocampus, and midbrain relative to Coke delivery preceded by a circle of light. As with the taste test, equivalent knowledge about Pepsi delivery had no such effect; indeed, no brain areas showed a significant difference to Pepsi delivered with versus without brand knowledge. The hippocampus and DLPFC have both been previously implicated in modifying behavior based on emotion and affect. The DLPFC is commonly implicated in aspects of cognitive control, including working memory (e.g., Watanabe, 1996). Lesions to the DLPFC are also known to result in depression (Davidson, 2002), which is hypothesized to result from a decreased ability to use positive affect to modify behavior (Mineka et al., 1998). It has been proposed that the DLPFC is necessary for employing affective information in biasing behavior (Watanabe, 1996; Davidson and Irwin, 1999). This is consistent with our findings: labeling Coke in taste and imaging tasks both biases behavior and recruits DLPFC activity. Furthermore, both of these effects are lost when compared with the semianonymous Pepsi tasks.

The hippocampus has also been implicated in processing affective information, but this association is tied to its role in the acquisition and recall of declarative memories (Eichenbaum, 2000; Markowitsch et al., 2003). The hippocampus is especially important in relating information to "discontiguous" sensory cues, as in the case of our task in which there is a 10 s period (catch trials) between the cue and soda delivery (McEchron and Disterhoft, 1999; Christian and Thompson, 2003). Imaging data indicate that the hippocampus is also im-

portant in recalling affect-related information (Iidaka et al., 2003; Markowitsch et al., 2003). Our finding supports these data and suggests that the hippocampus may participate in recalling cultural information that biases preference judgments.

Importantly, the hippocampus and DLPFC are only two of several brain areas that have been implicated in biasing behavior based on affect. Other areas include the amygdala, ventral striatum, anterior cingulate cortex, posterior cingulate cortex, and orbitofrontal cortex (for example, Greene et al., 2001; Schall et al., 2002; Ochsner et al., 2002; Sanfey et al., 2003). To our knowledge, the experiments linking these brain areas to judgments all involve subjects making decisions through expressed motor behavior. In our experiment, by contrast, the scanning session involved only the percept of soda, with no instruction to make preference decisions. It is an interesting possibility that the hippocampus and DLPFC are specifically involved in biasing perception based on prior affective bias, whereas the other brain areas listed above are more involved in altering behavioral output.

Determining preferences in our experiment appears to result from the interaction of two separate brain systems situated principally in the prefrontal cortex. The ventro-medial region of the prefrontal cortex plays a prominent role when preferences are determined solely from sensory information. The relative activity in the VMPFC is a very good indicator of which sensory stimulus is preferred by the subject. However, cultural influences have a strong influence on expressed behavioral preferences. We found this to be particularly the case with Coca-Cola, for which brand information significantly influences subjects' expressed preferences. We hypothesize that cultural information biases preference decisions through the dorsolateral region of the prefrontal cortex, with the hippocampus engaged to recall the associated information. These two systems appear to function independently in our experiment, since VMPFC activity was unaffected by brand knowledge. In judging stimuli based on multifaceted sensory and cultural influences, independent brain systems appear to cooperate to bias preferences.

Experimental Procedures

A total of 67 subjects participated in the study (38 male, 29 female; aged 19–50 years, mean \pm SD: 28.0 \pm 7.6 years old). All subjects gave informed consent to participate in the study; the Baylor Institutional Review Board approved the experimental paradigm. Each subject participated in one of three similarly designed experiments (group 1: anonymous, $n_1 = 16$, one failed to finish fMRI scanning due to technical problems; group 2: anonymous, $n_2 = 17$; group 2: semianonymous Coke, $n_3 = 16$; group 3: semianonymous Pepsi, $n_4 = 18$). For all four groups, subjects were first given a taste test and then completed the fMRI study. Subjects were not instructed to abstain from drinking prior to the experiment, but all subjects reported that they enjoyed the drinks.

Taste Tests

Taste tests consisted of 3 rounds (groups 1, 3, and 4) or 15 rounds (group 2) of forced-choice preference decisions between two cups of cola drinks (10 mL each). Coke and Pepsi were decarbonated in both the taste tests and scanning experiments in order to ensure reliable delivery through the plastic tubes required for the scanning experiment. In each round of the taste tests, cups were presented in random order. In the anonymous test (groups 1 and 2), both cups

in each round were unlabeled; one cup of each pair contained Coke, while the other contained Pepsi (Figure 1A). In group 3 (semanymous Coke), one cup in each pair was labeled "Coke" and the other was unlabeled; both cups contained Coke. The semianonymous Pepsi experiment (group 4) was the same as the semianonymous Coke experiment except the cups both contained Pepsi and one of each pair of cups was labeled "Pepsi." During the test, the Coke and Pepsi bottles (2 liter versions) were explicitly visible on the table in front of the subject. For each pair, subjects selected the drink they preferred; we refer to this in the text as *behavioral preference*. Between each taste-pair, subjects waited at least 40 s and rinsed with equal volumes of bottled water.

Subjects who participated in the 3-trial anonymous (group 1 and group 3) taste test were recalled after a period of several months to perform the 15-trial anonymous taste test. The number of selections each subject made to Coke in the 3-trial and in the 15-trial version of the test are compared in Figure 1C.

fMRI Task

Subjects lay supine with their head in the scanner bore and viewed a back-projected computer-generated image via a 45° mirror. Subjects were instructed to watch the screen and swallow the colas when they were delivered; there was no other task to perform. Whole-brain echo planar images were acquired with BOLD contrast and a repetition time of 2 s (performed on a Siemens Allegra 3T scanner; echo time 40 ms, flip angle 90°).

We employed a paradigm described in previous work and discussed more fully in the Results section (Figure 1B, McClure et al., 2003). Subjects were given a series of training pairings in which cues predicted the squirt of cola at a fixed delay of 6 s. After 20 training pairings, six catch trials, where drink delivery was delayed by 4 s, were randomly interspersed in the last 25 pairings. This produced a total of 35 pairings for each cue, or a grand total of 70 squirts of cola. The light-drink pairings were separated by a random time delay ranging from 6 s to 16 s. We focused on the amplitude of evoked BOLD responses to Pepsi and Coke delivered at delayed (unexpected) times. The paradigm was completed in two scanning runs of 291 and 296 scans, respectively.

For the group 1 and 2 subjects, the light stimuli were a red circle and a yellow circle, with the association of light color to soda flavor randomized across subjects (i.e., red→Coke, yellow→Pepsi or red→Pepsi, yellow→Coke). For the semianonymous experiments, an image of a Coke can (group 3) and an image of a Pepsi can (group 4) were used to cue the occurrence of one of the sodas. A red or yellow circle (randomized across subjects) predicted the other. Both sodas were the same (group 3, both Coke; group 4, both Pepsi).

Drink Delivery in Scanner

Individual squirts of Coke and Pepsi (0.8 mL each) were delivered to subjects through cooled plastic tubes held in the subjects' mouths with plastic mouthpieces. The volume of soda delivered on each squirt was sufficient to allow the subjects to fully taste the soda but were small enough to allow them to easily swallow while lying in the scanner. A computer-controlled syringe pump (Harvard Apparatus, Holliston, MA) allowed for precise delivery of the colas.

Image Analysis

Analysis of functional imaging data was performed using SPM2 (Friston et al., 1995). Raw data were realigned, corrected for slice-timing artifacts, normalized to a canonical spatial axis (resampled at 4 mm × 4 mm × 4 mm), and smoothed with an 8 mm Gaussian kernel. In the general linear model analysis, individual events were assumed to evoke known changes in hemodynamic response through time. Linear regression of these assumed response dynamics to measured BOLD signals produced amplitude estimates for each class of effect in the experiment. Student's t tests were performed over these amplitude estimates, constituting a random effects analysis over the subjects. Regions are reported as active that contain a minimum of three contiguous voxels each significant at $p < 0.001$. To protect against type I errors, we further ensured that these regions were composed of at least seven voxels at $p < 0.005$ (Forman et al., 1995). Both of these thresholds were further tested

and ensured to satisfy a false discovery rate of $q < 0.05$ for all of our result images (Genovese et al., 2002).

The motion parameters derived from image realignment were used to ensure that head movement did not corrupt our results. This analysis was performed in two separate ways: (1) mean event-related movement parameters were calculated and tested against the null hypothesis of no significant movement at any time point, and (2) the movement parameters were entered into a general linear model and regressed to the MRI data in the regions of interest (ROI). Neither of these efforts revealed any significant effect.

Acknowledgments

We thank Ann Harvey, Kristen Pfeiffer, and Nathan Apple for their help conducting this experiment. Further thanks go to Leigh Nystrom for his consultation on data analysis. This work was supported by the Kane Family Foundation (PRM) and NIDA grant DA-11723 (P.R.M.). All trademarks appearing in this article are the property of their respective owners.

Received: February 2, 2004

Revised: July 22, 2004

Accepted: September 2, 2004

Published: October 13, 2004

References

- Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'Connor, E., and Breiter, H.C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron* 32, 537–551.
- Anderson, A.K., and Sobel, N. (2003). Dissociating intensity from valence as sensory inputs to emotion. *Neuron* 39, 581–583.
- Armony, J.L., and Dolan, R.J. (2001). Modulation of auditory neural responses by a visual context in human fear conditioning. *Neuroreport* 12, 3407–3411.
- Bechara, A., Damasio, A.R., Damasio, H., and Anderson, S.W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* 50, 7–15.
- Berns, G.S., McClure, S.M., Pagnoni, G., and Montague, P.R. (2001). Predictability modulates human brain response to reward. *J. Neurosci.* 21, 2793–2798.
- Chacko, E., McDuff, I., and Jackson, R. (2003). Replacing sugar-based soft drinks with sugar-free alternatives could slow the progress of the obesity epidemic: have your Coke and drink it too. *N. Z. Med. J.* 116, U649.
- Christian, K.M., and Thompson, R.F. (2003). Neural substrates of eyeblink conditioning: acquisition and retention. *Learn. Mem.* 10, 427–455.
- Davidson, R.J. (2002). Anxiety and affective style: role of prefrontal cortex and amygdala. *Biol. Psychiatry* 51, 68–80.
- Davidson, R.J., and Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends Cogn. Sci.* 3, 11–21.
- Dolan, R.J., De Gleder, B., and Morris, J.S. (2001). Cross-modal binding of fear in voice and face. *Proc. Natl. Acad. Sci. USA* 98, 10006–10010.
- Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. *Nat. Neurosci.* 1, 41–50.
- Erk, S., Spitzer, M., Wunderlich, A.P., Galley, L., and Walter, H. (2002). Cultural objects modulate reward circuitry. *Neuroreport* 13, 2499–2503.
- Ford, E.S., Mokdad, A.H., and Giles, W.H. (2003). Trends in waist circumference among U.S. adults. *Obes. Res.* 11, 1223–1231.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., and Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647.
- Friston, K.J., Holmes, A.P., Worsley, K., Poline, J.P., Frith, C.D., and Frackowiak, R.S.J. (1995). Statistical parametric maps in functional brain imaging: a general linear approach. *Hum. Brain Mapp.* 2, 189–210.

- Genovese, C.R., Lazar, N.A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878.
- Gottfried, J.A., and Dolan, R.J. (2003). The nose smells what the eye sees: crossmodal visual facilitation of human olfactory perception. *Neuron* 39, 375–386.
- Gottfried, J.A., O'Doherty, J., and Dlova, R.J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science* 301, 1104–1107.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., and Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.
- Iidaka, T., Terashima, S., Yamashita, K., Okada, T., Sadato, N., and Yonekura, Y. (2003). Dissociable neural responses in the hippocampus to the retrieval of facial identity and emotion: an event-related fMRI study. *Hippocampus* 13, 429–436.
- Knutson, B., Adams, C.M., Fong, G.W., and Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.
- Laurienti, P.J., Burdette, J.H., Wallace, M.T., Yen, Y.-F., Field, A.S., and Stein, B.E. (2002). Activity in visual and auditory cortex can be modulated by influences from multiple senses. *J. Cogn. Neurosci.* 14, 420–429.
- Laurienti, P.J., Wallace, M.T., Maldjian, J.A., Susi, C.M., Stein, B.E., and Burdette, J.H. (2003). Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Hum. Brain Mapp.* 19, 213–223.
- McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38, 339–346.
- McEchron, M.D., and Disterhoft, J.F. (1999). Hippocampal encoding of non-spatial trace conditioning. *Hippocampus* 9, 385–396.
- Markowitzsch, H.J., Vandekerckhove, M.M., Lanfermann, H., and Russ, M.O. (2003). Engagement of lateral and medial prefrontal areas in the ecphory of sad and happy autobiographical memories. *Cortex* 39, 643–665.
- Mineka, S., Watson, D., and Clark, L.A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annu. Rev. Psychol.* 49, 377–412.
- Ochsner, K.N., Bunge, S.A., Gross, J.J., and Gabrieli, J.D. (2002). Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229.
- O'Doherty, J.P., Dayan, P., Friston, K.J., Critchley, H., and Dolan, R.J. (2003a). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.
- O'Doherty, J.P., Critchley, H., Deichmann, R., and Dolan, R.J. (2003b). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J. Neurosci.* 23, 7931–7939.
- Popkin, B.M., and Nielsen, S.J. (2003). The sweetening of the world's diet. *Obes. Res.* 11, 1325–1332.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758.
- Schall, J.D., Stuphorn, V., and Brown, J.W. (2002). Monitoring and control of action by the frontal lobes. *Neuron* 36, 309–322.
- Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., and Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* 39, 701–711.
- Stein, B.E., London, N., Wilkinson, L.K., and Price, D.D. (1996). Enhancement of perceived visual intensity by auditory stimuli: A psychophysical analysis. *J. Cogn. Neurosci.* 8, 497–506.
- Stein, B.E., Wallace, M.T., and Stanford, T.R. (1999). Development of multisensory integration: transforming sensory input into motor output. *Ment. Retard. Dev. Disabil. Res. Rev.* 5, 72–85.
- Wallace, M.T., and Stein, B.E. (1997). Development of multi-sensory integration in cat superior colliculus. *J. Neurosci.* 17, 2429–2444.
- Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature* 382, 629–632.
- Wyatt, H.R. (2003). The prevalence of obesity. *Prim. Care* 30, 267–279.
- Zimmet, P. (2003). The burden of type 2 diabetes: are we doing enough? *Diabetes. Metab.* 29, 6S9–18.
- Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappelow, J.C., and Berns, G.S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron* 42, 509–517.

Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related fMRI study

Florin Dolcos,^{a,b,*} Kevin S. LaBar,^a and Roberto Cabeza^a

^aCenter for Cognitive Neuroscience, Duke University, Durham, NC 27708-0999, USA

^bCentre for Neuroscience, University of Alberta, Edmonton, AB, Canada, T6G 2S2

Received 3 November 2003; revised 9 April 2004; accepted 18 May 2004

Prefrontal cortex (PFC) activity associated with emotional evaluation and subsequent memory was investigated with event-related functional MRI (fMRI). Participants were scanned while rating the pleasantness of emotionally positive, negative, and neutral pictures, and memory for the pictures was tested after scanning. Emotional evaluation was measured by comparing activity during the picture rating task relative to baseline, and successful encoding was measured by comparing activity for subsequently remembered versus forgotten pictures (Dm effect). The effect of arousal on these measures was indicated by greater activity for both positive and negative pictures than for neutral ones, and the effect of valence was indicated by differences in activity between positive and negative pictures. The study yielded three main results. First, consistent with the valence hypothesis, specific regions in left dorsolateral PFC were more activated for positive than for negative picture evaluation, whereas regions in right ventrolateral PFC showed the converse pattern. Second, dorsomedial PFC activity was sensitive to emotional arousal, whereas ventromedial PFC activity was sensitive to positive valence, consistent with evidence linking these regions, respectively, to emotional processing and self-awareness or appetitive behavior. Finally, successful encoding (Dm) activity in left ventrolateral and dorsolateral PFC was greater for arousing than for neutral pictures. This finding suggests that the enhancing effect of emotion on memory formation is partly due to an augmentation of PFC-mediated strategic, semantic, and working memory operations. These results underscore the critical role of PFC in emotional evaluation and memory, and disentangle the effects of arousal and valence across PFC regions associated with different cognitive functions.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Hemispheric asymmetry; Frontal lobes; Affect; Declarative memory; Semantic encoding; Self-referential processing; Neuroimaging

Introduction

The domain of cognitive neuroscience of emotion has grown dramatically during the last decade. As a result of this develop-

ment, various effects of emotion on brain activation associated with different perceptual and cognitive functions have been revealed (reviewed in Davidson and Irwin, 1999; Lane and Nadel, 2000; Phan et al., 2002). To understand these diverse effects, researchers have often divided emotion into its basic underlying constructs. One dimensional approach to emotion emphasizes the contribution of two orthogonal components, namely *arousal* and *valence* (Lang et al., 1993; Russell, 1980). *Arousal* refers to a continuum that varies from calm to excitement, whereas *valence* refers to a continuum that varies from positive to negative with neutral in the middle (for methods to assess these two dimensions, see Bradley and Lang, 1994). The vast majority of studies has focused on the limbic system and particularly on the amygdala, whereas other components of the emotional processing network, such as the prefrontal cortex (PFC), have received relatively less attention. Recent studies of amygdala function have attempted to tease apart the relative contributions of the aforementioned affective dimensions to task performance (e.g., Anderson et al., 2003; Hamann et al., 2002; Phelps and Anderson, 1997; Phelps et al., 1998). However, the contribution of these factors to emotional processing in other frontolimbic regions is not well understood, and the available evidence is contradictory. To address this imbalance, the present functional MRI (fMRI) study focused on the role of PFC regions in emotional processing.

In particular, we investigated the effects of arousal and valence on emotional evaluation and emotional memory. Emotional evaluation refers to the perception and categorization of emotional stimuli, and emotional memory refers to the modulatory effect of emotion on different stages of memory processing, including encoding, consolidation, and retrieval. In the domain of emotional evaluation, the amygdala is assumed to be involved in the rapid detection of the basic emotional properties of incoming stimuli, whereas PFC is assumed to be involved in higher-order emotional evaluation processes, which operate in close interaction with other cognitive functions and with behavioral goals (Davidson and Irwin, 1999). In the domain of emotional memory, the existing studies have focused on the amygdala and identified arousal-mediated effects at encoding that predict subsequent memory (Cahill et al., 1996; Canli et al., 2000, 2002; Dolcos et al., 2003, 2004; Hamann et al., 1999), but there is little understanding of the contribution of other brain regions, such as PFC regions. Although it is assumed

* Corresponding author. Brain Imaging and Analysis Center, Room 163, Bell Building, PO Box 3918, Duke University, Durham, NC 27710. Fax: +1-919-681-7033.

E-mail address: fdolcos@duke.edu (F. Dolcos).

Available online on ScienceDirect (www.sciencedirect.com.)

that activation during emotional evaluation tasks plays a role in memory, few studies have explicitly examined the relationship between emotional evaluation and memory. Thus, the overarching goal of the present study was to carefully investigate the contribution of the PFC to arousal and valence effects on emotional evaluation and memory.

Different PFC subregions are likely to make distinct contributions to emotional evaluation, but information about this issue is scarce. A basic anatomical distinction in this domain is between lateral and medial PFC regions. According to one prevailing view, the role of lateral PFC regions in emotional evaluation is primarily related to valence. The valence hypothesis states that the left PFC is dominant in the processing of positive emotions, whereas the right PFC is dominant in the processing of negative emotions (Davidson, 1995; Davidson and Irwin, 1999). This hypothesis is inspired by evidence from lesion literature and is mainly supported by electrophysiological evidence from EEG recordings. Neuropsychological evidence shows that patients with left hemisphere lesions tend to experience negative emotions, such as sadness (Morris et al., 1996; Paradiso et al., 1999a), whereas patients with right hemisphere damage are biased toward experiencing positive emotions, such as euphoria (e.g., Starkstein et al., 1989). The results of some electrophysiological studies are consistent with the valence hypothesis and support the idea that this valence-related PFC lateralization may depend either on transiently induced affective states or on stable personality traits (Aftanas et al., 2001; Davidson, 1995; Davidson and Irwin, 1999; Tomarken et al., 1992; Wheeler et al., 1993).

However, electrophysiological studies do not provide an accurate localization of the sources of these valence effects, and, overall, the evidence supporting the valence hypothesis has been mixed. First, neither lesion nor electrophysiological evidence has always been consistent with the valence hypothesis (e.g., Borod, 1992; Borod et al., 1998; Dolcos and Cabeza, 2002; Hagemann et al., 1998). Second, functional neuroimaging evidence is also inconclusive. Whereas some studies support the valence hypothesis (e.g., Canli et al., 1998), some studies do not report valence-related hemispheric asymmetry in PFC (e.g., Baker et al., 1997; George et al., 1995; Lane et al., 1997a,b,c; Pardo et al., 1993; Teasdale et al., 1999). One possible reason why the results have been mixed is that arousal and valence are often not distinguished carefully (but see Canli et al., 1998); hence, valence effects might have been confounded with arousal effects. Thus, the first goal of the present study was to investigate the valence hypothesis and identify the specific PFC regions involved, in conditions where positive and negative stimuli were matched in arousal and other potentially confounding factors were controlled.

As for the role of medial PFC regions in emotional evaluation, different hypotheses have been suggested. For instance, orbitofrontal areas of medial PFC have been linked to the rewarding nature of stimuli (e.g., O'Doherty et al., 2001; Rolls, 2000), and anteromedial areas have been related to more personal and subjective aspects of experiencing internal states (e.g., Frith and Frith, 1999). Although medial PFC regions have been strongly associated with emotional processing, it is unclear whether the role of these regions is related to arousal or to valence. Given that medial PFC regions are systematically activated by emotional stimuli, regardless of their valence (for a review, see Phan et al., 2002), PFC involvement could be attributed to its role in the processing of arousal. This notion is consistent with evidence supporting medial PFC involvement in processing emotionally arousing stimuli

irrespective of valence (Lane et al., 1997a,b,c; Reiman, 1997; Reiman et al., 1997; Schneider et al., 1995; Teasdale et al., 1999). On the other hand, there is also evidence suggesting valence-related specificity in medial PFC (e.g., George et al., 1995; Paradiso et al., 1999b). In particular, medial PFC has been associated with affiliative behaviors and appetitive or reward circuits (e.g., Rolls, 2000). A recent metaanalysis of functional neuroimaging studies of emotion (Wager et al., 2003) found that, overall, medial PFC activity was associated with approach or appetitive tasks. To address this issue, the second goal of the present study was to determine whether the role of medial PFC in emotional processing is primarily related to arousal or to valence, or whether there are subregions within medial PFC differently involved in arousal and valence.

Turning to emotional memory, the most basic phenomenon to explain in this domain is why arousing events (both positive and negative) are better remembered than neutral events (Bradley et al., 1992; Christianson, 1992). This effect has been attributed to the modulatory effect of the amygdala on the medial temporal lobe (MTL) memory system (McGaugh et al., 2002), and this modulation hypothesis has been confirmed by functional neuroimaging studies (Cahill et al., 1996; Canli et al., 2000, 2002; Dolcos et al., 2003, 2004; Hamann et al., 1999; Kilpatrick and Cahill, 2003). For example, we investigated this hypothesis using event-related fMRI and the subsequent memory paradigm (Paller and Wagner, 2002). In this paradigm, memory performance on a subsequent memory task is used to sort encoding items into two categories: remembered versus forgotten. Greater encoding activity for remembered than forgotten items, sometimes known as "the Dm (difference in memory) effect", is assumed to reflect successful encoding operations. Consistent with the modulation hypothesis, we found that the Dm effects in the amygdala and the MTL memory regions were greater for emotionally arousing pictures than for neutral pictures, and that the two Dm effects were more strongly correlated in the case of arousing pictures than in the case of neutral pictures (Dolcos et al., 2004).

Although this evidence strongly links the memory-enhancing effect of emotion to an MTL mechanism, it does not exclude the possibility that other brain regions, such as PFC, also play a major role. In fact, in functional neuroimaging studies, PFC regions are as strongly associated with successful encoding operations as MTL regions (e.g., Brewer et al., 1998; Paller and Wagner, 2002; Wagner et al., 1998). Also, the effects of several factors affecting encoding success, such as organizational strategies and attention, have been found to be mediated by changes in PFC activity (Anderson et al., 2000; Fletcher et al., 1998; Kensinger et al., 2003). Moreover, studies using transcranial magnetic stimulation (TMS) have shown that PFC activity is actually necessary for successful encoding (Epstein et al., 2002; Grafman and Wassermann, 1999; Rossi et al., 2001). Thus, it is quite likely that the enhancing effect of emotion on encoding is mediated not only by MTL but also by PFC. Yet, very little is known about the role of PFC on emotional memory formation (see, however, Canli et al., 2002; Kilpatrick and Cahill, 2003). For example, it is uncertain if the Dm effect in PFC is enhanced by emotion, similar to what we found in MTL (Dolcos et al., 2004). It is also unclear whether this putative effect is due to arousal or valence, and which specific PFC regions are involved. Thus, the third goal of the study was to investigate the role of PFC in the formation of emotional memory.

The method we employed has two main features: (1) it distinguishes between activity associated with emotional evalua-

tion and emotional memory, and (2) it distinguishes between the effects of arousal and valence. Participants were scanned while rating the pleasantness of arousing pictures (positive and negative) and non-arousing pictures (neutral), and after scanning, they recalled the contents of the pictures. Stimuli were selected from a standardized set of pictures that allows experimental control over arousal and valence characteristics (Lang et al., 1997), which has been largely used in neuroimaging studies of emotion (e.g., Dolan et al., 2000; Hamann et al., 1999, 2002; Lane et al., 1997a,b,c, 1999; Liberzon et al., 2000; Paradiso et al., 1999b; Taylor et al., 1998, 2000). Evaluation activity was measured by comparing activity during picture rating to the baseline activity, and successful encoding activity was measured by comparing activity for subsequently remembered versus subsequently forgotten pictures (Dm effect). Given that positive and negative pictures were both more arousing than neutral pictures, arousal effects should affect both positive and negative pictures. Given that positive and negative pictures were matched in arousal, differences between them should reflect valence effects rather than arousal effects. Thus, the effect of arousal was defined as greater activity for both positive and negative pictures than for neutral pictures, and the effect of valence, as differences between activity for positive and negative pictures.

To summarize, we investigated three main issues. First, we investigated the valence hypothesis, and in particular, what specific left and right PFC subregions would be sensitive to valence effects. Second, we investigated the role of medial PFC in emotional evaluation, and specifically, whether activity in this region is primarily sensitive to arousal or valence, or whether subregions can be distinguished. Finally, we investigated the role of PFC in emotional memory, and particularly, the relationship of Dm effects with stimulus arousal and valence, and their localization within PFC.

Methods

Subjects

Sixteen young (25 + 4.6 years), right-handed, healthy women participated in the study. Female participants were chosen because evidence suggests that they are more likely to display strong physiological responses to emotional stimuli (Lang et al., 1993) and report more intense emotional experiences (Shields, 1991) than men. All participants consented to a protocol approved by Duke University Institutional Review Board.

Materials

Stimuli consisted of 60 positive, 60 negative, and 60 neutral pictures selected from the International Affective Picture System (IAPS) picture database (Lang et al., 1997), on the basis of their normative arousal and valence scores. The mean arousal scores (1 = calm, 9 = excited) were 6.0 for positive ($SD = 2.2$), 6.15 for negative ($SD = 2.2$), and 3.15 for neutral pictures ($SD = 2.0$). Thus, positive and negative pictures had similar high arousal scores, whereas neutral pictures had low arousal scores. The mean valence scores (1 = negative, 5 = neutral, 9 = positive) were 7.1 for positive ($SD = 1.7$), 2.3 for negative ($SD = 1.5$), and 5.2 for neutral ($SD = 1.4$). To equate the emotional and neutral categories for visual complexity and content (e.g., human presence), the IAPS pictures were supplemented with neutral pictures from other sources (Yama-

saki et al., 2002). Also, given the evidence that dorsal–ventral PFC regions are differently involved in the processing of visual stimuli depending on their spatial content (e.g., Goldman-Rakic, 1995), it was important to determine that spatial/scene content did not differ between emotional and neutral stimuli. To investigate this idea, we asked 10 participants to rate our stimuli using a scene content scale (1 = no scene, 4 = very high scene content). The ratings for emotional (1.86) and neutral (1.93) pictures were similar ($T = 0.64$, $P > 0.5$), suggesting an equivalent amount of scene/spatial information across the stimulus categories.

Experimental design

The pool of 180 pictures was divided into six sets of 30 pictures (10 positive, 10 negative, and 10 neutral), which were randomly assigned to six study blocks. Six different block orders were randomly assigned to the participants. To avoid the induction of long-lasting mood states, the pictures within each block were pseudo-randomized so that no more than two pictures of the same valence were consecutively presented. Functional MR images were recorded while subjects viewed emotional and neutral pictures. The pictures were presented, using an LCD projector, to a screen located behind the subjects' crown that subjects could see via an angled mirror. Each picture was presented for 3 s and followed by a 12-s fixation cross. Participants were instructed to experience any feelings or thoughts the pictures might elicit in them, and to rate each picture in a 3-point pleasantness scale (1 = negative, 2 = neutral, 3 = positive). Rating the emotional valence of stimuli was employed because paying attention to emotional responses elicited by various stimuli is associated with deep encoding, which results in better subsequent memory performance. This task also provides an estimation of subjects' emotional responses. Nothing was mentioned about a subsequent memory test, before or during the encoding task, and hence learning was incidental. Incidental learning was preferred because intentional learning may interfere with the experience of emotions, and because differences in voluntary attention may complicate the interpretation of subsequent memory effects.

Forty-five minutes after the scanning session, subjects performed an unexpected cued-recall test conducted outside the MRI suite. Subjects were provided with one- or two-word written cues for each picture (e.g., snake, building, skydivers), and had to describe in writing, and in as much detail as they could, the pictures that they remembered. Similar to the procedure employed in our previous ERP study (Dolcos and Cabeza, 2002), participants were asked to provide enough relevant details (e.g., about the number of elements, color, action, etc.) so that an outsider could identify each picture and discriminate it from similar studied pictures (e.g., a brown snake facing viewer vs. several small green snakes). The test lasted until participants could not recall any additional pictures or until a maximum of 50 min had elapsed. Two raters were involved in scoring participants' responses, and only those pictures whose description was detailed enough to allow both identification and discrimination were classified as remembered.

MRI data acquisition

Anatomical scanning

Neuroimaging was performed using a 1.5 T GE scanner. A T1-weighted sagittal localizer series was first acquired. The anterior (AC) and posterior commissures (PC) were identified in the

midsagittal slice, and 34 contiguous oblique slices were prescribed parallel to the AC-PC plane. High-resolution T1-weighted structural images were acquired with a 450-ms TR (repetition time), a 9-ms TE (echo time), a 24-cm FOV (field of view), a 256^2 matrix, and a slice thickness of 3.75 mm. A second series of 46 oblique T1-weighted images perpendicular to the AC-PC was then acquired using the same imaging parameters.

Functional scanning

Thirty-four contiguous gradient-echo echoplanar images (EPIs) sensitive to blood oxygen level dependent (BOLD) contrast were acquired parallel to the AC-PC plane, using the same slice prescription described above for the near-axial structural images. The EPIs were acquired with a 3-s TR, 40-ms TE, one radio frequency excitation, 24-cm FOV, 64^2 image matrix, and a 90° flip angle. Slice thickness was 3.75 mm, resulting in cubic 3.75 mm³ isotropic voxels.

fMRI data analysis

Image preprocessing

Image preprocessing and statistical analyses were performed using SPM99 (<http://www.fil.ion.ucl.ac.uk/spm/>). Functional images were corrected for acquisition order and realigned to correct for motion artifacts. Anatomical images were coregistered with the first functional images for each subject, and then both anatomical and functional images were spatially normalized to a standard stereotactic space, using the Montreal Neurological Institute (MNI) templates implemented in SPM99. Subsequently, functional images were spatially smoothed using an 8-mm isotropic Gaussian kernel.

Statistical analyses

Statistical analyses were separately performed to assess both emotion-related (emotional evaluation) and memory-related differences (emotional encoding) between emotional and neutral pictures. The images were defined as unpleasant, neutral, or pleasant based on the IAPS ratings. The use of the IAPS score was justified by two reasons. First, different from the rating scores, which due to technical limitations during scanning (i.e., the response box had only three response options) did not allow fine evaluations of the subjects' emotional response, the IAPS scores are based on more sophisticated methods of assessing emotional arousal and valence, and allow much finer dissociations. Second, although some subjects classified some of the images differently than the norms, the high correlation between the average picture scores as rated by the subjects and the IAPS valence scores ($R = 0.9$, $P < 0.0001$) suggest that participants' classification was highly consistent with the normative data. Therefore, the latter scores were used to dissociate the effect of arousal and valence on brain activity associated with emotional evaluation and emotional memory formation. For each subject, task-related activity was identified by convolving a vector of the onset times of the stimuli with a synthetic hemodynamic response (HDR) and its temporal derivative. The general linear model, as implemented in SPM99, was used to model the effects of interest and other confounding effects (e.g., session effects and magnetic field drift). Functional images were proportionally scaled to the whole-brain signal.

Group analyses were conducted using random-effects models to assess the effect of arousal and valence on emotional evaluation and emotional encoding. In the present manuscript, we report the PFC

results, with a focus on lateral and medial cortices, excepting motor and cingulate regions. The MTL results were reported in different manuscripts (Dolcos et al., 2003, 2004). Conjunction analyses were used to identify brain regions more activated in two conditions (e.g., positive and negative) than in a third condition (e.g., neutral). This was done using the ImCalc feature in SPM, and according to the following formula: [(Condition 1 T score > 2.01) \times (Condition 2 T score > 2.01)]. This procedure yields a mask containing only those voxels that were significantly activated above $T = 2.01$ ($P = 0.0316$) in each and both contrasts. The probability of finding a voxel that is independently significant in each and both contrasts (i.e., the joint probability) can be estimated by multiplying the probabilities for each contrast: $0.0316 \times 0.0316 = P < 0.001$ (e.g., Allan et al., 2000; Cabeza et al., 2002).

The resulting conjunction masks provided information about the extent of the overlapping activations associated with the conditions involved in the conjunction, but not about the intensity of the overlapping activity. For this, conjunction T maps were calculated by multiplying the T values for the conditions of interest in the overlapping regions. The conjunction T maps were calculated according to the following formula: [(Condition 1 T score) \times (Condition 2 T score) \times (Conjunction mask of Condition 1 and Condition 2)]. These are the T values we report.

For emotional evaluation, the effect of arousal was measured as greater activity (compared to baseline) for arousing stimuli (positive and negative pictures) than for non-arousing stimuli (neutral pictures). This was done by identifying regions that showed both (1) greater activity for positive than for neutral pictures, and (2) greater activity for negative than for neutral pictures. That is, the *effect of arousal on emotional evaluation* was defined as [(positive $>$ neutral) conj (negative $>$ neutral)]. The effect of valence was measured as significant differences between positive and negative pictures. Since the valence scales had neutral as an intermediate value between positive and negative, we further required that regions associated with positive or negative valence had to be more activated in these conditions than in the neutral condition. That is, the *effect of positive valence on emotional evaluation* was defined as [(positive $>$ negative) conj (positive $>$ neutral)] and the *effect of negative valence on emotional evaluation* was defined as [(negative $>$ positive) conj (negative $>$ neutral)].

The same kind of analyses were performed for emotional encoding, except that instead of using activity during picture rating compared to baseline, we used differences in activity between remembered and forgotten items (D_m = activity for remembered pictures – activity for forgotten pictures). D_m activity was separately calculated for each picture category (e.g., D_m positive = positive remembered – positive forgotten), and then the effects of arousal and valence were identified by comparing the three types of D_m activity. The *effect of arousal on emotional memory* was defined as [(D_m positive $>$ D_m neutral) conj (D_m negative $>$ D_m neutral)]. The *effect of positive valence on emotional memory* was defined as [(D_m positive $>$ D_m negative) conj (D_m positive $>$ D_m neutral)] and the *effect of negative valence on emotional memory* was defined as [(D_m negative $>$ D_m positive) conj (D_m negative $>$ D_m neutral)]. To make sure that the differences between D_m s occurred due to positive activations in the condition of interest and were not driven by deactivations in the other conditions, the conjunction maps were inclusively masked with the activation maps showing the main effect of memory (D_m) for the condition of interest at $P < 0.05$. For instance, for the latter comparison (i.e.,

$[(Dm \text{ negative} > Dm \text{ positive}) \text{ conj } (Dm \text{ negative} > Dm \text{ neutral})]$, the resulting map was masked with the activation map for Dm negative. Thus, the final conjunction map contained only the voxels that showed a significant Dm for negative pictures.

The bar graphs of fMRI activations were examined by extracting the mean effect size from the peak voxel of each region, as identified by the SPM conjunction analyses for each condition of interest and subject. The data extraction was accomplished using SPM99. The xyz coordinates provided by SPM, which are in Montreal Neurological Institute (MNI) brain space, were converted to xyz coordinates in Talairach and Tournoux's brain space (Talairach and Tournoux, 1988).

Results

Behavioral results

Valence ratings

The average valence scores (1 = negative, 2 = neutral, 3 = positive) as rated by the participants in the scanner were 1.14 ($SD = 0.16$) for negative pictures, 2.18 ($SD = 0.40$) for neutral pictures, and 2.64 ($SD = 0.26$) for positive pictures. All pairwise comparisons were significant ($P < 0.0001$). Thus, the subjects' rating scores were consistent with those provided in the IAPS norms (Lang et al., 1997). Further validating this consistency, the correlation between our subjects' average scores and the normed IAPS scores of the pictures used in the present study was highly significant ($R = 0.90$, $P < 0.0001$).

Memory performance

Arousing pictures, both positive and negative, were better recalled than neutral pictures. Out of 60 pictures per category, participants recalled an average of 52% positive, 53% negative, and 38% neutral pictures (SDs were 4.5, 4.8, and 4.8, respectively). An ANOVA yielded a significant picture type effect ($F(2,15) = 41.21$, $P < 0.0001$), and post-hoc contrasts showed that recall of positive and negative pictures was similar ($P > 0.05$), and greater than recall of neutral pictures ($P < 0.0001$).

fMRI results

The analyses on activity associated with emotional evaluation (rating-baseline) yielded dissociable PFC regions showing effects of arousal and valence. In lateral PFC, the main goal was to test the valence hypothesis. Consistent with this hypothesis, a valence-related hemispheric asymmetry was found: a left dorsolateral PFC region (BA 8/9; $xyz = -41, 21, 48$; $T = 8.63$) showed an effect of positive valence, whereas a right ventrolateral PFC region (BA 47; $xyz = 49, 33, -2$; $T = 17.03$) showed an effect of negative valence. As illustrated by Fig. 1, the left PFC region was more activated for positive than for negative pictures, whereas the right PFC region was more activated for negative than for positive pictures. However, this asymmetry was not present in the entire lateral PFC: other dorsolateral PFC areas (BA 9) showed a bilateral ($xyz = 49, 9, 24/-38, 9, 24$; $T = 20.02/T = 15.41$) effect of negative valence. The effect of arousal in lateral PFC was evident in an area of the right inferior frontal gyrus (BA 47; $xyz = 30, 18, -17$; $T = 9.57$).

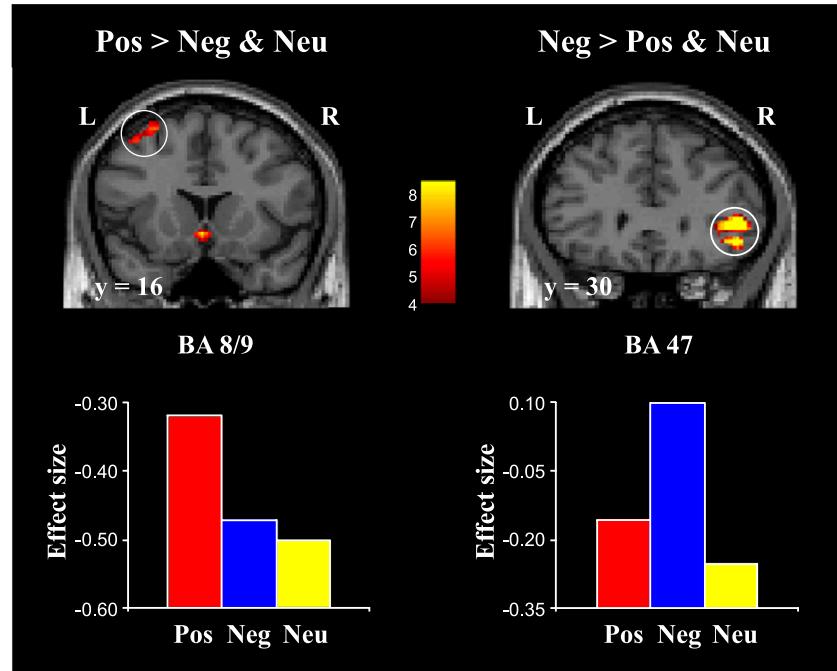


Fig. 1. Activity in lateral PFC showed evidence for the valence hypothesis. In the left hemisphere, a dorsolateral PFC region (BA 8/9) was more activated during evaluation of positive pictures than during evaluation of negative pictures, whereas in the right hemisphere, ventrolateral regions (BA 47) were more activated for negative than for positive pictures. The upper panels show the activation maps overlaid on high-resolution coronal anatomical images, and the bar graphs at the bottom show the effect size as extracted from the peak-voxels identified in the conjunction analyses (see Methods). The numbers at the left-bottom corner of the upper panels (e.g., $y = 16$) indicate the coordinate in MNI space. The color bar located between the upper panels indicates the conjunction T values. L = left, R = right; Pos = positive, Neg = negative, Neu = neutral; BA = Brodmann Area.

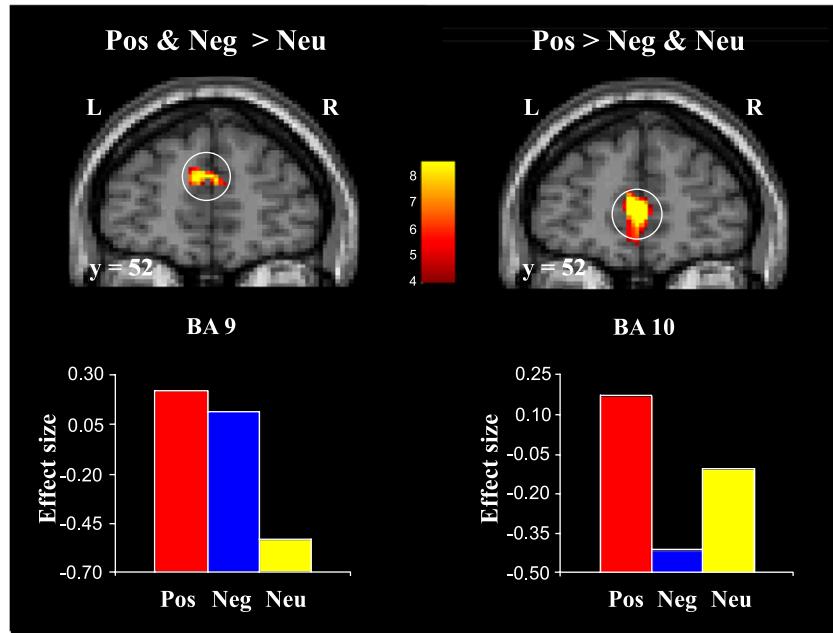


Fig. 2. Activity in medial PFC identified dissociable regions associated with arousal and positive valence. Dorsomedial PFC (BA 9) activity was sensitive to arousal (Pos and Neg > Neu), whereas ventromedial PFC (BA 10) activity was sensitive to positive valence (Pos > Neg and Neu). L = left, R = right; Pos = positive, Neg = negative, Neu = neutral; BA = Brodmann Area.

In medial PFC, a dorsal–ventral distinction was found during emotional evaluation in which arousal effects were in a dorsal region (BA 9; $xyz = -4, 52, 19$; $T = 11.45$), whereas valence effects were found in more ventral locations—orbitofrontal and antero-medial cortices (BA 10; $xyz = 0, 58, -10/0, 58, 4$; $T = 14.72/T = 30.2$). As illustrated by Fig. 2, the dorsomedial PFC region was

more activated during evaluation of both pleasant and unpleasant pictures than during evaluation of neutral pictures, whereas the ventral regions were more activated during the evaluation of positive pictures.

The analysis of activity associated with emotional memory ($Dm = \text{remembered} - \text{forgotten}$) also yielded PFC regions

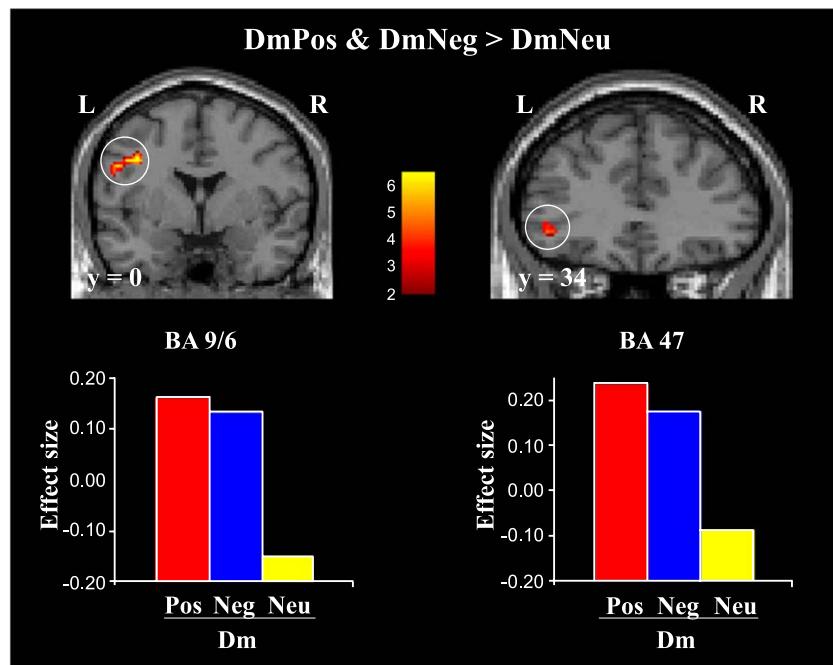


Fig. 3. Arousal enhanced successful encoding activity (Dm) in left PFC. Compared to the Dm for neutral pictures, the Dm for arousing pictures (both positive and negative) was greater in left ventrolateral (BA 47) and dorsolateral (BA 9/6) PFC regions. L = left, R = right; Pos = positive, Neg = negative, Neu = neutral; BA = Brodmann Area; $Dm = \text{remembered} - \text{forgotten}$.

showing effects of arousal and valence. The areas showing effects of arousal were found only in the lateral PFC, and included both ventral (BA 47; $xyz = -49, 29, -1$; $T = 5.48$) and dorsal (BA 9/6; $xyz = -38, 2, 31$; $T = 8.64$) locations. As illustrated by Fig. 3, in these lateral PFC regions, the Dm for positive pictures and the Dm for negative pictures were both greater than the Dm for neutral pictures. The average effect size values for remembered and forgotten pleasant, unpleasant, and neutral items, respectively, are as follows: BA 47 ($-0.22, -0.46, -0.13, -0.30, -0.31$, and -0.22) and BA 9/6 ($-0.24, -0.41, -0.09, -0.22, -0.31$, and -0.16). The areas showing effects of valence included medial (BA 9; $xyz = -11, 38, 29$; $T = 6.41$) and lateral (BA 45; $xyz = -49, 22, 3$; $T = 8.51$) locations, both showing an effect of positive valence.

Discussion

The present study yielded three main findings relevant for understanding PFC contributions to emotional evaluation and memory. First, consistent with the valence hypothesis, during emotional evaluation, specific left dorsolateral PFC areas showed greater activity for positive than for negative pictures, whereas right ventrolateral PFC areas showed the converse pattern. Second, also during emotional evaluation, dorsomedial PFC activity was sensitive to arousal (greater activation for both positive and negative pictures relative to neutral ones), whereas ventromedial PFC activity was sensitive to valence (greater activation for positive pictures relative to negative ones). Finally, demonstrating the role of lateral PFC in emotional memory, arousal enhanced successful encoding activity in left ventrolateral and dorsolateral PFC. These results provide evidence for multiple, regionally specific emotional influences on PFC function. The implications of the findings are discussed in separate sections below.

Lateral PFC showed a hemispheric asymmetry consistent with the valence hypothesis

As illustrated by Fig. 1, a left PFC region was more activated for positive than for negative pictures and a right PFC region showed the converse pattern. This finding is consistent with other evidence supporting the valence hypothesis (Aftanas et al., 2001; Canli et al., 1998; Davidson, 1995; Tomarken et al., 1992; Wheeler et al., 1993), but it extends this evidence in two ways.

First, the present finding demonstrates hemispheric asymmetry effects predicted by the valence hypothesis under conditions in which positive and negative stimuli were matched for arousal and visual properties, and in which the effects of valence could be distinguished from the effects of arousal on a trial-by-trial basis. As noted before (see Methods), positive and negative stimuli had similar normative scores of arousal and were equivalent in terms of complexity, presence of human figures, and other lower-level visual features. Participant ratings obtained on-line during encoding confirmed the valence manipulation and were highly correlated with normative valence scores. Moreover, the inclusion of neutral pictures allowed us to disentangle valence effects from arousal effects. Since arousal effects are by definition common to positive and negative pictures, these effects should appear as greater activity for both positive and negative pictures than for neutral pictures. As illustrated by Fig. 1, this was not the case in either the left or the right PFC regions comprising valence-related hemispheric asymmetry. Additionally, in contrast to previous findings

(Canli et al., 1998), which could not determine whether hemispheric asymmetries were stimulus-specific or reflected sustained changes in affective states, the present study provides evidence supporting the valence hypothesis using an event-related design in which stimuli were randomized during scanning.

Second, the present results not only demonstrated a valence-related hemispheric asymmetry but also identified the specific left and right PFC regions associated with positive and negative valence. The valence hypothesis has been primarily supported using electrophysiological methods (Aftanas et al., 2001; Davidson, 1995; Tomarken et al., 1992; Wheeler et al., 1993), which do not allow a good localization of neural sources. Evidence for the valence hypothesis was also found using fMRI (Canli et al., 1998) but using analyses that collapsed activity over a whole hemisphere. In contrast, the present result shows that the left PFC region specifically sensitive to positive valence is in dorsolateral cortex (middle frontal gyrus; BA 8/9) whereas the right PFC region particularly sensitive to negative valence is in ventrolateral cortex (inferior frontal gyrus; BA 47). A more inferior sector of BA9, in contrast, did not show such asymmetry effects, and was sensitive to negative valence bilaterally.

What are the implications of this region-specific hemispheric asymmetry in the processing of emotional valence? As described above, lesion, electrophysiological, as well as the available neuroimaging evidence supporting the valence hypothesis lacks the regional specificity necessary to link activity in the regions identified here with possible differential involvement in processing emotional valence. One possibility is to broadly explain the role of these regions based on the available neuroimaging evidence concerning their involvement in various tasks. Another possibility is to link our findings with more specific evidence concerning dorsal–ventral dissociations in the lateral PFC. As concerning the former, neuroimaging evidence has associated the dorsolateral sectors of PFC, including the BAs 8 and 9 with a variety of tasks, involving perceptual, attentional, imagistic, and mnemonic operations, but typically they have been associated with working memory tasks (see Cabeza and Nyberg, 2000, for a review). In addition, portions of BA 8 are thought to be part of the so-called frontal eye field, although the overlap between this region in human and nonhuman primates is still controversial (Koyoma et al., 2004). Interestingly, related to its role in emotional processing, BA 8 has been identified in tasks involving rating the pleasantness of facial stimuli (Nakamura et al., 1998). As concerning the role of ventrolateral PFC regions, including BA 47, their function has been associated with semantic memory operations, as well as with interference control and inhibitory processes (Miller and Cohen, 2001; Smith and Jonides, 1999; see also Cabeza and Nyberg, 2000). Specifically related to the involvement in emotional processing, right ventrolateral PFC has been implicated in the inhibition of negative emotions (Petrovic et al., 2002).

Turning to the evidence concerning the asymmetry in the dorsal–ventral dimension, two main views have been identified regarding the role of lateral PFC. According to one view, dorsolateral PFC regions are more involved in manipulating working memory contents whereas ventrolateral PFC regions are more involved in simple maintenance operations (D’Esposito et al., 2000; Owen et al., 1999; Petrides, 1995). Another view (Davidson and Irwin, 1999) specifically relates the function of lateral PFC to the role of emotion in guiding and organizing behavior in a motivationally consistent manner (Frijda, 1988). According to this view, dorsolateral PFC is involved in a particular form of working

memory—*affective working memory*—responsible for the representation of goal-related emotional states, whereas the ventrolateral sector of the PFC is involved in the simple representation of elementary emotional states (Davidson and Irwin, 1999).

Combining these ideas, one may speculate that the left dorsolateral PFC activity for positive stimuli reflects the maintenance and manipulation of positive information in working memory during the valence-rating task, whereas the right ventrolateral PFC activity reflects the inhibition (avoidance) of negative information. Of course, these ideas are ad hoc and require independent confirmation. For instance, since other dorsolateral PFC regions were associated with evaluation of negative pictures, it would be important to clarify that working memory-related activity in the left PFC region specifically associated with positive pictures is related to the maintenance of appetitive goals (e.g., detection and processing of positive stimuli).

Dissociable regions of medial PFC were associated with arousal and positive valence

As illustrated by Fig. 2, dorsomedial PFC activity was sensitive to arousal, whereas ventromedial PFC activity was sensitive to positive valence. Previous evidence is consistent with the idea of possible segregation among medial PFC regions with respect to their involvement in emotional processing, but it is unclear whether the role of these regions is related to arousal or to valence. On the one hand, a metaanalysis of neuroimaging studies of emotion suggests that medial PFC regions are systematically activated by emotional stimuli, regardless of their valence (for a review, see Phan et al., 2002). This finding suggests a nonspecific involvement of medial PFC in emotional processing, probably mediated by arousal, and is consistent with some neuroimaging studies reporting dorsomedial activations associated with the processing of emotional stimuli, regardless of their valence (Lane et al., 1997a,b,c; Reiman, 1997; Reiman et al., 1997; Teasdale et al., 1999). On the other hand, another metaanalysis (Wager et al., 2003) found that, overall, medial PFC activity was associated with approach or appetitive tasks. Although approach-withdrawal and valence dimensions are not identical, they do overlap, and thus this finding suggests possible valence-related specificity in medial PFC function (see also George et al., 1995; Paradiso et al., 1999b). However, comparisons across studies are complicated by differences in stimuli, methods, and participants.

In the present study, we demonstrated a dorsal–ventral dissociation in the medial PFC, within-subjects and under controlled conditions. By identifying specific medial PFC regions sensitive to arousal and valence, this finding complements and reconciles previous functional neuroimaging evidence suggesting nonspecific (Lane et al., 1997a,b,c; Reiman, 1997; Reiman et al., 1997; Teasdale et al., 1999) versus valence-specific involvement of medial PFC during emotional processing (George et al., 1995; Paradiso et al., 1999b). This finding is also consistent with the results of an ERP study (Dolcos and Cabeza, 2002) where we found arousal versus valence dissociations at midline frontal electrodes. Although spatial resolution of ERP did not allow us to separate these effects topographically, it allowed dissociations in timing: there was a faster effect (500–800 ms) of positive valence (positive > negative = neutral) and a delayed effect (after 800 ms) of arousal (positive = negative > neutral). Since the valence-related ERP effect occurred in an earlier time-window (see also Cuthbert et al., 2000; Dillon et al., submitted), we interpreted it as reflecting

a bias toward detecting and processing positive stimuli. Given the fact that, overall, the medial PFC in the present study was more activated for positive than for negative pictures, this interpretation may also apply to the current finding.

An alternative account is that the sensitivity of medial PFC to positive valence reflected greater self-engagement in the processing of positive pictures compared to negative and neutral pictures. A number of recent neuroimaging studies associated medial PFC activity with self-referential processing (e.g., Cabeza et al., in press; Frith and Frith, 1999; Gusnard et al., 2001; Kelley et al., 2002). For instance, in a review of the literature, Frith and Frith (1999) suggested that activity in ventral medial PFC was specifically associated with emotional aspects of self-processing. It should be noted, however, that the appetitive and self-engagement accounts of the present medial PFC activation are compatible. For example, participants could have been more likely to relate the pictures to their own self and life in the case of positive pictures than in the case of negative and neutral pictures. The medial PFC region is frequently activated in functional neuroimaging studies of autobiographical memories. For example, we found this region to be more activated during the recognition of photographs taken by oneself than during the recognition of photographs taken by others (Cabeza et al., in press).

Arousal enhanced successful encoding activity in left PFC

Emotional arousal enhanced successful encoding (Dm) activity in lateral PFC. As illustrated by Fig. 3, compared to the Dm for neutral pictures, the Dm for arousing pictures was greater in left ventrolateral and dorsolateral PFC. These findings suggest that the enhancing effect of emotion on memory formation (i) is partly mediated by changes in PFC activity, (ii) is mainly related to arousal, and (iii) may involve an amplification of semantic processing and working memory operations mediated by lateral PFC regions.

Research on the neural bases of the enhancing effect of emotion on memory formation has emphasized the role of the amygdala and its interactions with MTL memory regions (Cahill et al., 1996; Canli et al., 2000, 2002; Dolcos et al., 2003, 2004; Hamann et al., 1999; Kilpatrick and Cahill, 2003; McGaugh et al., 2002). The present results expand this line of research by showing that the enhancing effect of emotion on memory formation is also mediated by changes in PFC activity. However, the effects of emotion on MTL and PFC are likely to enhance different memory mechanisms. Given the functions typically attributed to these regions (Moscovitch, 1992; Simons and Spiers, 2003), it is reasonable to assume that in MTL, emotion enhances the storage and consolidation of memory representations, whereas in PFC, it enhances strategic encoding processes.

A second implication of the present findings is that the enhancing effect of emotion on memory formation is primarily related to arousal rather than to valence. In our MTL study (Dolcos et al., 2003, 2004), we also found that the Dm increase was related to arousal rather than to valence. Thus, although valence-related Dm increases also occur, it seems fair to conclude that arousal is the main factor modulating the neural mechanisms of memory formation. This conclusion is consistent with our behavioral results, which showed that compared to non-arousing neutral pictures, memory is better for arousing positive and arousing negative pictures, with no significant difference between these two conditions. Thus, from the point of view of memory, a negative event can be as effective as a positive event (see also Talarico et al., in press).

The specific PFC regions where the Dm was increased by arousal suggest that arousing events are better remembered because they receive deeper semantic processing and working memory resources during encoding. The Dm was enhanced by arousal in left ventrolateral (BA 47) and dorsolateral (BA 9/6) PFC regions (see Fig. 3). The left ventrolateral region is an area that many functional neuroimaging studies have associated with encoding processes (for a review, see Cabeza and Nyberg, 2000), including event-related fMRI studies using the subsequent memory paradigm (Brewer et al., 1998; Kirchhoff et al., 2000; Paller and Wagner, 2002; Wagner et al., 1998). Since the role of this region in encoding is generally attributed to semantic processing (Kapur et al., 1996; Poldrack et al., 1999; Shallice et al., 1994), the present finding suggests that arousal facilitated successful encoding by increasing semantic processing of the information in the pictures. It is possible that arousal also enhanced perceptual encoding processes mediated by right PFC (Brewer et al., 1998). However, since picture memory was tested using verbal recall, the effects of arousal on perceptual encoding were probably not detected. To detect such effects, it would be necessary to test memory with a nonverbal task, such as picture recognition (e.g., Brewer et al., 1998).

Finally, the effect of arousal on left dorsolateral PFC is likely to reflect the augmentation of the working memory processes typically associated with this region (D'Esposito et al., 2000; Owen et al., 1999; Petrides, 1995). Thus, it is possible that the contents of arousing events not only receive deeper semantic processing but are also maintained longer or manipulated more intensely in working memory, leading to better retention. It should be noted that this interpretation is compatible with the idea of dorsolateral PFC involvement during emotional evaluation of positive pictures, since the specific regions involved during positive evaluation versus successful encoding of arousing pictures are slightly different. It is possible that in one case, the involvement of working memory operations is related to the maintenance of appetitive-goals, whereas in the latter, the maintenance or manipulation of emotionally arousing information leads to better subsequent memory.

Conclusions

Using an fMRI paradigm that distinguished between activity related to emotional evaluation and emotional memory and between the effects of arousal and valence, the present study yielded three main results. First, during emotional evaluation, PFC activity showed a hemispheric asymmetry consistent with the valence hypothesis. A left dorsolateral PFC region was sensitive to positive valence, possibly reflecting the maintenance of positive information in working memory, whereas a right ventrolateral PFC region was sensitive to negative valence, possibly reflecting the inhibition of negative information. Second, dorsomedial PFC activity was sensitive to arousal, whereas ventromedial PFC activity was sensitive to positive valence, possibly reflecting the involvement of these regions in general processing of emotional information (dorsomedial PFC), and self-awareness or appetitive behavior (ventromedial PFC). Finally, successful encoding activity was enhanced by arousal in left ventrolateral and dorsolateral PFC regions, possibly reflecting an enhancement of strategic, semantic, and working memory operations. Although further research is required, these findings strongly suggest that different PFC regions are sensitive to arousal

and to valence, and that they play an important role in the evaluation of emotional stimuli and in processes that lead to better memory for emotional events.

Acknowledgments

We thank David Beckmann for assistance with data analysis. This study was supported by NIH grants R01 AG19731 (RC) and R01 DA14094 (KSL), and a NARSAD Young Investigator Award (KSL). FD was supported by a Chia PhD Scholarship and a Dissertation Fellowship from the University of Alberta (Canada), and a Research Assistantship from Duke University (USA).

References

- Aftanas, L., Varlamov, A., Pavlov, S., Makhnev, V., Reva, N., 2001. Event-related synchronization and desynchronization during affective processing: emergence of valence-related time-dependent hemispheric asymmetries in theta and upper alpha band. *Int. J. Neurosci.* 110 (3–4), 197–219.
- Allan, K., Dolan, R.J., Fletcher, P.C., Rugg, M.D., 2000. The role of the right anterior prefrontal cortex in episodic retrieval. *NeuroImage* 11 (3), 217–227.
- Anderson, N.D., Iidaka, T., Cabeza, R., Kapur, S., McIntosh, A.R., Craik, F.I., 2000. The effects of divided attention on encoding- and retrieval-related brain activity: a PET study of younger and older adults. *J. Cogn. Neurosci.* 12 (5), 775–792.
- Anderson, A.K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D.G., Glover, G., Gabrieli, J.D., Sobel, N., 2003. Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* 6 (2), 196–202.
- Baker, S.C., Frith, C.D., Dolan, R.J., 1997. The interaction between mood and cognitive function studied with PET. *Psychol. Med.* 27 (3), 565–578.
- Borod, J.C., 1992. Interhemispheric and intrahemispheric control of emotion: a focus on unilateral brain damage. *J. Consult. Clin. Psychol.* 60, 339–348.
- Borod, J.C., Obler, L.K., Erhan, H.M., Grunwald, I.S., Cicero, B.A., Welkowitz, J., Santschi, C., Agosti, R.M., Whalen, J.R., 1998. Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology* 12 (3), 446–458.
- Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25 (1), 49–59.
- Bradley, M.M., Greenwald, M.K., Petry, M.C., Lang, P.J., 1992. Remembering pictures: pleasure and arousal in memory. *J. Exp. Psychol., Learn., Mem., Cogn.* 18 (2), 379–390.
- Brewer, J.B., Zhao, Z., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1998. Making memories: brain activity that predicts how well visual experience will be remembered. *Science* 281, 1185–1187.
- Cabeza, R., Nyberg, L., 2000. Imaging cognition: II. An empirical review of 275 PET and fMRI studies. *J. Cogn. Neurosci.* 12 (1), 1–47.
- Cabeza, R., Dolcos, F., Graham, R., Nyberg, L., 2002. Similarities and differences in the neural correlates of episodic memory retrieval and working memory. *NeuroImage* 16 (2), 317–330.
- Cabeza, R., Prince, S.E., Daselaar, S.M., Greenberg, D.L., Budde, M., Dolcos, F., LaBar, K.S., Rubin, D.C., 2004. Brain activity during episodic retrieval of autobiographical and laboratory events: an fMRI study using a novel photo paradigm. *J. Cogn. Neurosci.* (in press).
- Cahill, L., Haier, R.J., Fallon, J., Alkire, M.T., Tang, C., Keator, D., Wu, J., McGaugh, J.L., 1996. Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proc. Natl. Acad. Sci. U. S. A.* 93, 8016–8021.

- Canli, T., Desmond, J.E., Zhao, Z., Glover, G., Gabrieli, J.D., 1998. Hemispheric asymmetry for emotional stimuli detected with fMRI. *NeuroReport* 9 (14), 3233–3239.
- Canli, T., Zhao, Z., Brewer, J., Gabrieli, J.D.E., Cahill, L., 2000. Event-related activation in the human amygdala associated with later memory for individual emotional experience. *J. Neurosci.* 20 (RC99), 1–5.
- Canli, T., Desmond, J.E., Zhao, Z., Gabrieli, J.D.E., 2002. Sex differences in the neural basis of emotional memories. *Proc. Natl. Acad. Sci. U. S. A.* 99 (16), 10789–10794.
- Christianson, S.-A., 1992. *The Handbook of Emotion and Memory: Research and Theory*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J., 2000. Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol. Psychol.* 52, 95–111.
- Davidson, R.J., 1995. Cerebral asymmetry, emotion and affective style. In: Davidson, R.J., Hugdahl, K. (Eds.), *Brain Asymmetry*. MIT Press, Cambridge, MA, pp. 361–387.
- Davidson, R.J., Irwin, W., 1999. The functional neuroanatomy of emotion and affective style. *Trends Cogn. Sci.* 3 (1), 11–20.
- D'Esposito, M., Postle, B.R., Rypma, B., 2000. Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Exp. Brain Res.* 133 (3–11).
- Dillon, D.G., Cooper, J.J., Grent-t-Jong, T., Woldorff, M.G., LaBar, K.S. Dissociation of event-related potentials indexing arousal, valence, and semantic cohesion during emotional stimulus encoding, submitted.
- Dolan, R.J., Lane, L., Chua, P., Fletcher, P., 2000. Dissociable temporal lobe activations during emotional episodic memory retrieval. *NeuroImage* 11, 203–209.
- Dolcos, F., Cabeza, R., 2002. Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral Pictures. *Cogn. Affect Behav. Neurosci.* 2 (3), 252–263.
- Dolcos, F., Graham, R., LaBar, K., Cabeza, R., 2003. Coactivation of the amygdala and hippocampus predicts better recall for emotional than for neutral pictures. *Brain and Cognition* 51, 221–223.
- Dolcos, F., LaBar, K.S., Cabeza, R., 2004. Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42, 855–863.
- Epstein, C.M., Sekino, M., Yamaguchi, K., Kamiya, S., Ueno, S., 2002. Asymmetries of prefrontal cortex in human episodic memory: effects of transcranial magnetic stimulation on learning abstract patterns. *Neurosci. Lett.* 320 (1–2), 5–8.
- Fletcher, P.C., Shallice, T., Dolan, R.J., 1998. The functional roles of prefrontal cortex in episodic memory: I. Encoding. *Brain* 121, 1239–1248.
- Frijda, N.H., 1988. The laws of emotion. *Am. Psychol.* 43 (5), 349–358.
- Frith, C.D., Frith, U., 1999. Interacting minds—A biological basis. *Science* 286 (5445), 1692–1695.
- George, M.S., Ketter, T.A., Parekh, P.I., Horwitz, B., Herscovitch, P., Post, R.M., 1995. Brain activity during transient sadness and happiness in healthy women. *Am. J. Psychiatry* 152 (3), 341–351.
- Goldman-Rakic, P.S., 1995. Architecture of the prefrontal cortex and the central executive. *Ann. N. Y. Acad. Sci.* 769, 71–83.
- Grafman, J., Wassermann, E., 1999. Transcranial magnetic stimulation can measure and modulate learning and memory. *Neuropsychologia* 37, 159–167.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.* 98 (7), 4259–4264.
- Hagemann, D., Naumann, E., Becker, G., Maier, S., Bartussek, D., 1998. Frontal brain asymmetry and affective style: a conceptual replication. *Psychophysiology* 35 (4), 372–388.
- Hamann, S.B., Ely, T.D., Grafton, S.T., Kilts, C.D., 1999. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nat. Neurosci.* 2 (3), 289–293.
- Hamann, S.B., Ely, T.D., Hoffman, J.M., Kilts, C.D., 2002. Ecstasy and agony: activation of the human amygdala in positive and negative emotion. *Psychol. Sci.* 13 (2), 135–141.
- Kapur, S., Tulving, E., Cabeza, R., McIntosh, A.R., Houle, S., Craik, F.I., 1996. The neural correlates of intentional learning of verbal materials: a PET study in humans. *Conit. Brain Res.* 4, 243–249.
- Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F., 2002. Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14 (5), 785–794.
- Kensinger, E.A., Clarke, R.J., Corkin, S., 2003. What neural correlates underlie successful encoding and retrieval? A functional magnetic resonance imaging study using a divided attention paradigm. *J. Neurosci.* 23 (6), 2407–2415.
- Kilpatrick, L., Cahill, L., 2003. Amygdala modulation of parahippocampal and frontal regions during emotionally influenced memory storage. *NeuroImage* 20 (4), 2091–2099.
- Kirchhoff, B.A., Wagner, A.D., Maril, A., Stern, C.E., 2000. Prefrontal–temporal circuitry for episodic encoding and subsequent memory. *J. Neurosci.* 20 (16), 6173–6180.
- Koyoma, M., Hasegawa, I., Osada, T., Adachi, Y., Nakahara, K., Miyashita, Y., 2004. Functional magnetic resonance imaging of macaque monkeys performing visually guided saccade tasks: comparison of cortical eye field with humans. *Neuron* 41 (5), 895–907.
- Lane, D.R., Nadel, L., 2000. *Cognitive Neuroscience of Emotion*. Oxford Univ. Press, Inc., New York.
- Lane, R.D., Fink, G.R., Chau, P.M., Dolan, R.J., 1997a. Neural activation during selective attention to subjective emotional responses. *NeuroReport* 8 (18), 3969–3972.
- Lane, R.D., Reiman, E.M., Ahern, G.L., Schwartz, G.E., Davidson, R.J., 1997b. Neuroanatomical correlates of happiness, sadness, and disgust. *Am. J. Psychiatry* 154 (7), 926–933.
- Lane, R.D., Reiman, E.M., Bradley, M.M., Lang, P.J., Ahern, G.L., Davidson, R.J., Schwartz, G.E., 1997c. Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia* 35 (11), 1437–1444.
- Lane, R.D., Chua, P.M., Dolan, R.J., 1999. Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia* 37 (9), 989–997.
- Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O., 1993. Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* 30 (3), 261–273.
- Lang, P.J., Bradley, M.M., Cuthberg, B.N., 1997. International Affective Picture System [Pictures]. NIMH Center for the Study of Emotion and Attention, Gainesville.
- Liberzon, I., Taylor, S.F., Fig, L.M., Decker, L.R., Koeppe, R.A., Minoshima, S., 2000. Limbic activation and psychophysiological responses to aversive visual stimuli. Interaction with cognitive task. *Neuropsychopharmacology* 23 (5), 508–516.
- McGaugh, J.L., McIntyre, C.K., Power, A.E., 2002. Amygdala modulation of memory consolidation: interaction with other brain systems. *Neurobiol. Learn. Mem.* 78 (3), 539–552.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 67–202.
- Morris, P.L., Robinson, R.G., Raphael, B., Hopwood, M.J., 1996. Lesion location and poststroke depression. *J. Neuropsychiatry Clin. Neurosci.* 8 (4), 399–403.
- Moscovitch, M., 1992. Memory and working-with-memory: a component process model based on modules and central systems. *J. Cogn. Neurosci.* 4, 257–267.
- Nakamura, K., Kawashima, R., Nagumo, S., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Kubota, K., Fukuda, H., Kojima, S., 1998. Neuroanatomical correlates of the assessment of facial attractiveness. *NeuroReport* 9 (4), 753–757.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4 (1), 95–102.
- Owen, A.M., Herrod, N.J., Menon, D.K., Clark, J.C., Downey, S.P., Carpenter, T.A., Minhas, P.S., Turkheimer, F.E., Williams, E.J., Robbins, T.W., Sahakian, B.J., Petrides, M., Pickard, J.D., 1999. Redefining the functional organization of working memory processes within human lateral prefrontal cortex. *Eur. J. Neurosci.* 11 (2), 567–574.

- Paller, K.A., Wagner, A.D., 2002. Observing the transformation of experience into memory. *Trends Cogn. Sci.* 6 (2), 93–102.
- Paradiso, S., Chemerinski, E., Yazici, K.M., Tartaro, A., Robinson, R.G., 1999a. Frontal lobe syndrome reassessed: comparison of patients with lateral or medial frontal brain damage. *J. Neurol., Neurosurg. Psychiatry* 67 (5), 664–667.
- Paradiso, S., Johnson, D.L., Andreasen, N.C., O’Leary, D.S., Watkins, G.L., Ponto, L.L., Hichwa, R.D., 1999b. Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. *Am. J. Psychiatry* 156 (10), 1618–1629.
- Pardo, J.V., Pardo, P.J., Raichle, M.E., 1993. Neural correlates of self-induced dysphoria. *Am. J. Psychiatry* 150 (5), 713–719.
- Petrides, M., 1995. Functional organization of the human frontal cortex for mnemonic processing. Evidence from neuroimaging studies. *Ann. N. Y. Acad. Sci.* 769, 85–96.
- Petrovic, P., Kalso, E., Petersson, K.M., Ingvar, M., 2002. Placebo and opioid analgesia—Imaging a shared neuronal network. *Science* 295, 1737–1740.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage* 16 (2), 331–348.
- Phelps, E.A., Anderson, A.K., 1997. Emotional memory: what does the amygdala do? *Curr. Biol.* 7 (5), R311–R314.
- Phelps, E.A., LaBar, K.S., Anderson, A.K., O’Connor, K.J., Fulbright, R.K., Spencer, D.D., 1998. Specifying the contributions of the human amygdala to emotional memory: a case study. *Neurocase* 4, 527–540.
- Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E., Glover, G.H., Gabrieli, J.D., 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *NeuroImage* 10, 15–35.
- Reiman, E.M., 1997. The application of positron emission tomography to the study of normal and pathologic emotions. *J. Clin. Psychiatry* 58 (16), 4–12.
- Reiman, E.M., Lane, R.D., Ahern, G.L., Schwartz, G.E., Davidson, R.J., Friston, K.J., Yun, L.S., Chen, K., 1997. Neuroanatomical correlates of externally and internally generated human emotion. *Am. J. Psychiatry* 154 (7), 918–925.
- Rolls, E.T., 2000. The orbitofrontal cortex and reward. *Cereb. Cortex* 10 (3), 284–294.
- Rossi, S., Cappa, S.F., Babiloni, C., Pasqualetti, P., Miniussi, C., Carducci, F., Babiloni, F., Rossini, P.M., 2001. Prefrontal [correction of Prefontal] cortex in long-term memory: an “interference” approach using magnetic stimulation. *Nat. Neurosci.* 4 (9), 948–952.
- Russell, J., 1980. A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178.
- Schneider, F., Gur, R.E., Mozley, L.H., Smith, R.J., Mozley, P.D., Censis, D.M., Alavi, A., Gur, R.C., 1995. Mood effects on limbic blood flow correlate with emotional self-rating: a PET study with oxygen-15 labeled water. *Psychiatry Res.* 61 (4), 265–283.
- Shallice, T., Fletcher, P., Frith, C.D., Grasby, P., Frackowiak, R.S., Dolan, R.J., 1994. Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* 368, 633–635.
- Shields, S., 1991. Gender in the psychology of emotion: a selective research review. In: Strongman, K. (Ed.), *International Review of Studies on Emotion*. Wiley, New York, pp. 227–245.
- Simons, J.S., Spiers, H.J., 2003. Prefrontal and medial temporal lobe interactions in long-term memory. *Nat. Rev., Neurosci.* 4 (8), 637–648.
- Smith, E., Jonides, J., 1999. Storage and executive processes in the frontal lobes. *Science* 283, 657–660.
- Starkstein, S.E., Robinson, R.G., Honig, M.A., Parikh, R.M., Joselyn, J., Price, T.R., 1989. Mood changes after right-hemisphere lesions. *Br. J. Psychiatry* 155, 79–85.
- Talairach, J., Tournoux, P., 1988. *A Co-Planar Stereotactic Atlas of the Human Brain*. Thieme, Stuttgart, Germany.
- Talarico, J.T., LaBar, K.S., Rubin, D.C., 2004. Emotional intensity predicts autobiographical memory experience. *Mem. Cogn.* (in press).
- Taylor, S.F., Liberzon, I., Fig, L.M., Decker, L.R., Minoshima, S., Koeppe, R.A., 1998. The effect of emotional content on visual recognition memory: a PET activation study. *NeuroImage* 8 (2), 188–197.
- Taylor, S.F., Liberzon, I., Koeppe, R.A., 2000. The effect of graded aversive stimuli on limbic and visual activation. *Neuropsychologia* 38 (10), 1415–1425.
- Teasdale, J.D., Howard, R.J., Cox, S.G., Ha, Y., Brammer, M.J., Williams, S.C., Checkley, S.A., 1999. Functional MRI study of the cognitive generation of affect. *Am. J. Psychiatry* 156 (2), 209–215.
- Tomarken, A.J., Davidson, R.J., Wheeler, R.E., Doss, R.C., 1992. Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *J. Pers. Soc. Psychol.* 62 (4), 676–687.
- Wager, T.D., Phan, K.L., Liberzon, I., Taylor, S.F., 2003. Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *NeuroImage* 19 (3), 513–531.
- Wagner, A.D., Schacter, D.L., Rotte, M., Koutstaal, W., Maril, A., Dale, A.M., Rosen, B.R., Buckner, R.L., 1998. Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 281, 1188–1191.
- Wheeler, R.E., Davidson, R.J., Tomarken, A.J., 1993. Frontal brain asymmetry and emotional reactivity: a biological substrate of affective style. *Psychophysiology* 30, 82–89.
- Yamasaki, H., LaBar, K.S., McCarthy, G., 2002. Dissociable prefrontal brain systems for attention and emotion. *Proc. Natl. Acad. Sci. U. S. A.* 99 (17), 11447–11451.

LEARNING & MEMORY

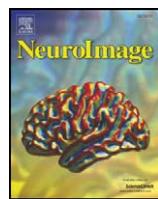


What is a memory? What is remembering, and can there be different kinds of recall? As soon as we think about memory, we probably think about the recall of vivid playbacks of what happened the other day, who said what, and how you felt.

Memory is a term covering a whole plethora of different functions and processes. Reflexes, instinctual behaviors, habits and reading are all in part memory processes, and with different kinds of access to prior learning. For example, you probably do not have a detailed movie-like playback of when you actually learned to ride a bike, but nevertheless you can easily jump on a bike and ride it more or less on auto pilot. Your conscious recollection of riding a bike is not necessary for you to use your learned and automatized behavior of riding a bike.

Recall can therefore also come in many shades. Think of a brand name. The first name that came to mind would be an instance of “free recall”. If I write a brand name like “Pepsi” you will recognize it, and “recognition” is a different memory process. If I ask you to judge your knowledge about the brand “BMW” you will access your semantic memory and make a self-evaluated rating of your own memory.

Not all memory types are relevant to consumer behavior – or are they? Maybe even the memory types we take most for granted are also part of our consumer choice process?



Tuning down the emotional brain: An fMRI study of the effects of cognitive load on the processing of affective images

Lotte F. Van Dillen ^{a,c,*}, Dirk J. Heslenfeld ^b, Sander L. Koole ^a

^a Department of Social Psychology, VU University, Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands

^b Department of Cognitive and Clinical Neuropsychology, VU University, Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands

^c Department of Social and Organizational Psychology, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands

ARTICLE INFO

Article history:

Received 1 August 2008

Revised 13 January 2009

Accepted 14 January 2009

Available online xxxx

Keywords:

Cognition

Emotion

Interaction

Amygdala

Insula

Dorsolateral frontal cortex

IAPS

ABSTRACT

The present research examines whether cognitive load can modulate the processing of negative emotional stimuli, even after negative stimuli have already activated emotional centers of the brain. In a functional magnetic resonance imaging (fMRI) study, participants viewed neutral and negative stimuli that were followed by an attention-demanding arithmetic task. As expected, exposure to negative stimuli led to increased activation in emotional regions (the amygdalae and the right insula). Subsequently induced task load led to increased activation in cognitive regions (right dorsolateral frontal cortex, right superior parietal cortex). Importantly, task load down-regulated the brain's response to negative stimuli in emotional regions. Task load also reduced subjectively experienced negative emotion in response to negative stimuli. Finally, coactivation analyses suggest that during task performance, activity in right dorsolateral frontal cortex was related to activity in the amygdalae and the right insula. Together, these findings indicate that cognitive load is capable of tuning down the emotional brain.

© 2009 Elsevier Inc. All rights reserved.

Negative emotional experiences are an inescapable aspect of human life. Although brief episodes of negative emotions can be adaptive (Öhman, 2007), negative emotions may become problematic when they persist over time. Indeed, there is mounting evidence that enduring negative emotional states impair both psychological and physical health (Lyubomirsky et al., 1999; Sapolsky, 1994). It is therefore vital for people to develop ways of effectively dealing with negative emotion.

One important way of dealing with negative emotion consists of minimizing the amount of attentional resources that are devoted to processing negative information (Ochsner and Gross, 2005; Nolen-Hoeksema et al., 1993; Van Dillen and Koole, 2007). More specifically, performing an attention-demanding task has been found to attenuate the emotional impact of negative stimuli (Erber and Tesser, 1992; Erthal et al., 2005; Glynn et al., 2002; Morrow and Nolen-Hoeksema, 1990; Pessoa et al., 2002; Van Dillen and Koole, 2007; in press). In recent years, the neural effects of task load on processing of emotional stimuli have begun to receive more systematic attention. Several studies have demonstrated that activity in emotion processing regions of the brain in response to negative emotional stimuli, such as the amygdalae, depend on the availability of attentional resources for processing of these stimuli (Blair et al., 2007; Erk et al., 2007; Okon-Singer et al., 2007; Mitchell et al., 2007; Pessoa et al., 2002; Van

Reekum et al., 2007). For example, in one study, negative visual distractors engaged the amygdalae during participants' judgements whether two bars were like oriented or not (Pessoa et al., 2002), but only when the difference in orientation of two bars was easy to judge. When the orientation of the two bars was difficult to judge, so that the central task became more attentionally demanding, the amygdalae no longer differentiated between the negative and neutral distractors. In another study (Erk et al., 2007), amygdalae responses to negative scenes were smaller when participants concurrently performed a working memory task that was highly rather than moderately demanding.

Although past work has made important progress, it remains unclear precisely how cognitive load influences the emotional brain. One possibility, which has been proposed in prior work, is that cognitive load prevents the processing of the emotional impact of negative stimuli altogether. From this perspective, cognitive load may cause an emotionally relevant stimulus to simply bypass emotional circuits. Given that most prior research relied on demanding distractor tasks that visually competed with emotional information (Erthal et al., 2005; Okon-Singer et al., 2007; Pessoa et al., 2002), cognitive load may have led participants to overlook some of the emotional information.

In support of this notion, Van Reekum et al. (2007) demonstrated that individuals spent less time fixating emotion-relevant stimulus features when they were instructed to reappraise negative visual scenes in less emotional terms than when they were instructed to merely attend to these pictures. This variation in gaze fixation could

* Corresponding author. Fax: +31 30 253 4718.

E-mail address: L.F.vanDillen@uu.nl (L.F. Van Dillen).

explain a significant amount of the reduction of activation in emotional brain regions during reappraisal compared to passive viewing. Accordingly, attentional deployment strategies (such as an additional task load or attending to non-emotional information) may allow people to manage unwanted emotions even before these emotions have been fully aroused (Gross, 2001).

Another possibility, which is highlighted in the present work, is that task load is capable of down-regulating emotional circuits even after these circuits have been mobilized. This would imply that emotional brain regions operate quite flexibly, in that they can still be modulated by contextual demands even when the initial emotional response has unfolded. In line with this, behavioral research suggests that task load can still modulate emotional processing when the task succeeds, rather than accompanies the emotional stimulus (Erber and Tesser, 1992; Gerin et al., 2006; Glynn et al., 2002; Van Dillen and Koole, 2007). For instance, participants reported less negative feelings in response to negative pictures when they subsequently tried to solve complex rather than simple math equations (Van Dillen and Koole, 2007). Perhaps then, performing a demanding task may similarly down-regulate the unfolding of the emotional brain response to a previously displayed emotional stimulus.

The present research was designed to investigate the neural dynamics by which cognitive load modulates the emotional brain response, even after this response has already been initiated. Importantly, in the present study, the cognitive task followed rather

than accompanied the display of the emotional stimulus. As a consequence, any effects of task load on neural responses to the emotional stimuli could not be accounted for by differences in visual processing strategies, such as controlling visual attention to emotional rather than to non-emotional stimulus features (Van Reekum et al., 2007).

In an experimental study, participants were exposed to neutral or negative emotional pictures, after which they performed an arithmetic task that made varying demands on processing resources (Ashcraft and Kirk, 2001), and rated their emotional state (Fig. 1). Throughout the experiment, participants' brain responses were monitored using functional magnetic resonance imaging (fMRI). We predicted that, relative to exposure to neutral pictures, exposure to negative pictures would activate both negative feelings and emotional brain circuits such as the amygdala and the insulae (Phan et al., 2002). In addition, we predicted that task load would increase activation in brain regions implicated in cognitive processing, such as the dorsolateral frontal cortex, the superior parietal cortex, and the dorsal occipital cortex (Duncan and Owen, 2000; Rypma et al., 1999).

Most importantly, we hypothesized that cognitive load would modulate subjectively reported emotional states and the activation of emotional brain circuits. That is, we predicted neural activity in regions implicated in emotional processing to decrease with increases in task load. Although the negative pictures should initiate a greater response in emotional brain areas than neutral pictures, we expected that the unfolding of this response over time would be attenuated when participants subsequently performed a highly demanding arithmetic task rather than a moderately demanding arithmetic task. In short, we predicted that task load would modulate the unfolding of the emotional brain response, even after this response was already initiated.

Method

Participants and design

Seventeen volunteers at the VU University Amsterdam (13 women, average age 20) took part in the experiment. All participants were right-handed and native Dutch speakers. The participants did not report any history of neurological or psychiatric problems. The ethical review board of the VU Medical Centre approved of the study and all volunteers provided written informed consent (according to the Declaration of Helsinki) after the study procedure had been explained to them. They were paid €20 for participation. The experimental design was a 2 (task load: high versus low)×2 (picture valence: neutral versus negative) factorial design, both factors within participants.

Procedure and equipment

Participants were invited to the lab to participate in a brain-imaging experiment. Before starting with the actual experiment, participants were instructed about the experimental set-up. Participants were then led to the scanner-room and positioned supine in the whole-body scanner, where they completed the actual experiment. All stimuli were back-projected onto a screen and viewed by participants through an angled mirror. The experiment consisted of a picture viewing task that contained four blocks of 32 experimental trials (128 trials in total). The order of the blocks was counterbalanced between participants and trials within blocks were displayed in random order. Each trial consisted of a picture followed by an arithmetic task and a mood scale. The pictures were selected from the International Affective Picture System (Lang et al., 2001). Based on published norms (ranging on a scale from 1 [most unpleasant] to 9 [most pleasant]), we selected two sets of 64 pictures, a negative set (pleasantness ratings lower than 2.50) and a neutral set (pleasantness

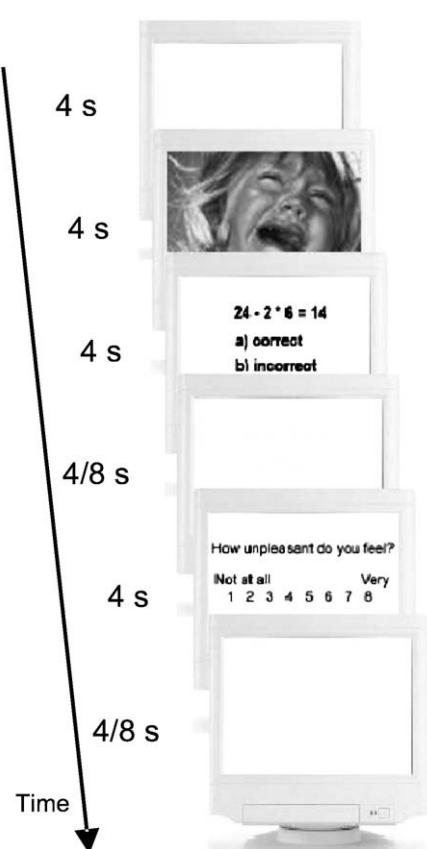


Fig. 1. Schematic depiction of a trial. After a 4 s blank display a picture was presented for 4 s and immediately followed by an arithmetic equation, that consisted of a product or a division combined with a summation or a subtraction, displayed for 4 s. Participants had to judge by a keyboard response whether the equation was correct or not. Pictures were either neutrally or negatively valenced; equations were either complex, such as '24-2*6=14', or simple, such as '5-2=3'. Following the equation, a blank display was presented for either 4 or 8 s, after which participants indicated on an eight-point scale, ranging from not at all (1) to very much (8), how unpleasant they felt at that moment. Participants had 4 s to make this response. Subsequently, a blank display was again presented for either 4 or 8 s and participants were instructed to relax.

ratings between 4.00 and 5.00). Negative pictures included images of scenes with burn victims, physical assaults, and angry faces. Neutral pictures depicted scenes of people in conversation, scenes of nature or buildings, and neutral faces. The two picture sets were matched as much as possible on dimensions such as living beings, faces, and complexity.

In each trial, a negative or neutral picture appeared on screen for 4 s. After picture presentation, participants had to perform an arithmetic task. The complexity of the task was randomly varied. In half of the trials, the arithmetic task consisted of a more complex equation such as '2*8+12=28'. These equations always combined a summation or subtraction with a product or division. In the remaining trials, the arithmetic task consisted of a much simpler equation, such as '7+2=9'. The latter equations only consisted of either a summation or a subtraction. Participants judged whether the equation was correct by pressing a button with either their left or right index finger. Participants had 4 s to make this response. For nine participants, the right button represented the correct response and the left button the incorrect response, while for eight participants this order was reversed.

At the end of each trial, participants rated, with a button response, how unpleasant they felt at that moment (from 1 = *not at all*, to 8 = *very much*). For eight participants, this scale was reversed (ranging from right to left). In between trials, participants were asked to relax. To avoid systematic overlap of BOLD responses within and between trials, the interval between the arithmetic task and the mood scale, as well as the interval between the mood scale and the beginning of the next trial, was set randomly to either 4,000 or 8,000 ms. The duration of each trial accordingly was 20, 24 or 28 s. The onset of each trial was synchronized to the onset of an fMRI volume.

Prior to the presentation of the first block, participants were given a block of 16 practice trials to get familiar with the experimental set-up and the scanner. After the experiment, participants were thanked for their efforts, debriefed, and paid by the experimenter.

A personal computer controlled presentation of the experimental trials and recorded participants' responses. The experimental trials were presented in E-prime (Psychology Software Tools, Inc., Pittsburgh, USA). Participants responded by pressing fiber-optic buttons (Lumitouch Photon Control, Burnaby, Canada).

MRI procedure and analysis

Brain imaging was performed on a 1.5 T Siemens Sonata scanner (Siemens Medical Systems, Erlangen, Germany) equipped with a volume head coil. Functional volumes consisted of 24 near axial slices acquired using an EPI sequence with the following parameters: repetition time = 2 s, echo time = 50 ms, flip angle = 90°, slice thickness = 4.2 mm, slice gap = 0.84 mm, acquisition matrix = 64 × 64 pixels, in-plane resolution = 3 × 3 mm. Series of 392 volumes were acquired in each of the four blocks of trials. Images were on-line motion corrected. After the functional session, a three-dimensional structural scan was acquired using a T1-weighted MP-RAGE sequence with the following scanning parameters: repetition time = 2730 ms, echo time = 3.43 ms, inversion time = 1000 ms, flip angle = 7°, 160 sagittal slices, slice thickness = 1 mm, acquisition matrix = 256 × 224 pixels, in-plane resolution = 1 × 1 mm.

Preprocessing and statistical analyses of the MRI data were performed using BrainVoyager 2000 software (Brain Innovation, Maastricht, The Netherlands). The first 2 volumes were discarded in order to avoid differences in T1 saturation. Voxel time-series of the remaining volumes were high-pass filtered (0.01 Hz), temporally smoothed (2.5 s FWHM Gaussian kernel), and corrected for slice acquisition times. Finally, volumes were 3D spatially smoothed (6 mm FWHM Gaussian kernel). Each functional run was manually co-registered to the individual 3D structural scan, re-sampled, and transformed into Talairach space (Goebel et al., 2001; Talairach and

Tournoux, 1988). Voxel time series were standardized and corrected for serial correlations.

fMRI data were first analyzed at each voxel (whole brain) and then specifically for a number of regions of interest (ROI's). The initial random-effects whole-brain analysis served to identify brain regions that responded in any way to negative pictures in our task. A multirun/multisubject GLM design matrix was constructed to model the relevant brain responses for each run and participant (Friston et al., 1999). The matrix consisted of regressors predicting hemodynamic responses to each combination of picture valence and task load, as well as the mood scale and possible arithmetic errors. ROI's were defined on the basis of the whole-brain activation obtained in response to negative pictures that were followed either by a simple or a complex arithmetic task. We used a "negative-fixation" contrast because we did not want to miss any region responding to negative pictures. We also included regressors for arithmetic errors and negative mood reports, as these may activate emotional brain regions as well, and may increase the error term if left unexplained. All regressors were convolved with a standard hemodynamic response function, and we analyzed the resulting beta weights at each voxel. The statistical threshold for this random-effects whole-brain analysis was $p = .05$ after Bonferroni correction for multiple comparisons.

The whole-brain analysis revealed activation clusters in a number of regions, such as the bilateral amygdalae, the bilateral inferior insulae, the right dorsolateral frontal cortex and the right superior parietal cortex (see Table 1 for all ROI's). Peaks of activation in each region were located, significant voxels surrounding those peaks were selected, and their time courses were averaged for a subsequent detailed analysis of the shape of the responses at each ROI.

For each ROI and subject, we computed the course of the hemodynamic responses to each valence/load combination by means of a deconvolution analysis. A GLM regressor was assigned to each of 8 fMRI volumes following the onset of a picture, separately for each combination of picture valence and task load. This allowed us to compute the shape of the hemodynamic response without prior assumptions, independent for each condition, ROI and subject. These deconvolved hemodynamic responses were then broken into time frames of 4 s duration (i.e., fMRI volume 1–2, 3–4, 5–6, 7–8), and used in subsequent random-effects analyses of the detailed interactions between task load and picture valence. The motivation for the 4 s (2 volumes) time frames was a compromise between temporal specificity and sufficient signal-to-noise. That is, we wanted to analyze the

Table 1
Statistical regions of interest

Brain region (Brodmann area, hemisphere)	Talairach coordinates (mm)			Volume (ml)
	x	y	z	
Dorsolateral frontal cortex (BA 6/44, right)	40	2	30	0.210
Superior parietal cortex (BA 7, right)	224	-59	53	0.331
Fasciculus uncinatus (BA 34, left)	-28	7	-11	0.353
Dorsal occipital cortex (left)	-22	-79	19	13.979
Dorsal occipital cortex (right)	23	-77	18	13.129
Ventral occipital cortex (left)	-22	-64	-5	30.286
Parahippocampal cortex (left)	-30	-25	-14	0.763
Parahippocampal cortex (right)	34	-19	-16	0.643
Medial anterior temporal cortex (BA 38, right)	34	-4	-23	0.221
Amygdala (left)	-19	-5	-9	0.211
Amygdala (right)	22	-8	-10	0.450
Inferior insula (left)	-31	-1	-9	465
Inferior insula (right)	31	-3	-9	0.173
Pulvinar (left)	-18	-24	0	0.118
Pulvinar (right)	19	-24	0	0.986
Locus coeruleus	0	-30	-28	0.244
Superior colliculus (left)	-6	-24	0	0.934
Superior colliculus (right)	6	-24	0	0.800

Brain areas responding to negative pictures as revealed by the random-effects whole-brain analysis ($p < .05$, corrected).

time courses of the effects of task load in detail, ideally at each volume, but then the number of tests would become large, and the signal-to-noise ratio of each test low. Therefore we grouped 2 volumes, which allowed us to analyze the time courses of the effects at sufficient temporal detail and with sufficient signal-to-noise.

Results

Arithmetic performance

To investigate whether our manipulation of task load was successful, we analyzed participants' correct responses and response times. Analyses of variance (ANOVA's) revealed that participants performed better on simple arithmetic tasks than on complex arithmetic tasks ($M=97\%$ correct, $SD=1.09$ versus $M=84\%$ correct, $SD=5.56$; $F(1, 16)=161.59$, $p<.0001$). Participants were also faster on simple than on complex arithmetic equations, $F(1, 16)=858.38$, $p<.0001$ ($M=1929$ ms, $SD=160$ versus $M=3083$ ms, $SD=145$). These effects confirm that the complex arithmetic equations were more difficult than the simple arithmetic equations. The effects of task complexity did not interact with picture valence for neither participants' correct responses ($F<1$) nor participants' response times ($F<1$).

Self-reported negative emotion

To examine whether task load modulates participants' self-reported negative emotion, we analyzed the effect of both task load (high, low) and picture valence (negative, neutral) on their self-reports, with participants response time differences between complex and simple arithmetic trials as a covariate. As expected, this ANCOVA yielded a significant effect of picture valence, $F(1, 16)=30.39$, $p=.000$. On average, participants reported more negative emotion after trials with negative pictures than after trials with neutral pictures (respectively, $M=4.55$, $SD=1.08$ versus $M=3.15$, $SD=1.14$).

The analysis further yielded the predicted interaction between task load and picture valence, $F(1, 16)=16.50$, $p=.001$. To interpret this effect, we analyzed the effects of task load separately in each valence condition. In line with previous research using the same paradigm (Van Dillen and Koole, 2007), there was no effect of task load in the neutral trials, $F<1$. By contrast, task load had a significant effect in the negative trials, $F(1, 16)=11.54$, $p=.004$. Participants reported less negative emotion when negative pictures were followed by a complex arithmetic equation rather than a simple arithmetic equation ($M=4.40$, $SD=1.07$ versus $M=4.70$, $SD=1.10$). Moreover, response time differences were unable to account for the effects of task load on negative moods, $F(1, 15)<1$. Thus, in line with previous findings (Van Dillen and Koole, 2007), task load modulated participants' self-reported negative emotion.

Brain regions involved in both emotion and task-related processes

To investigate whether task load modulated the unfolding of participants' emotional brain responses, we first identified brain regions that responded to negative emotional pictures in our task (see Method section). This yielded a number of areas, such as the bilateral amygdala, the bilateral inferior insulae, the right dorsolateral frontal cortex and the right superior parietal cortex. In Table 1, the regions of interest (ROI's) are given, with their xyz-coordinates and cluster sizes.

Subsequently, for each region and participant, we computed the course of the hemodynamic response to each picture/load combination by means of a deconvolution analysis, for the first 16 s following picture onset. That is, we estimated the shape of the hemodynamic response by assigning a GLM regressor to each of 8 fMRI volumes following picture onset. In order to assess the timing of the effects of our manipulations more precisely, the hemodynamic responses of the

selected regions were then broken into four time frames of two fMRI volumes each (1–4 s, 5–8 s, 9–12 s, and 13–16 s). These time frames were further tested in subsequent analyses of variance for effects of picture valence and task load.

These subsequent analyses revealed that there were no effects of picture valence or task load during the first time frame following picture onset ($t=1$ –4 s). Activation in response to the neutral and negative pictures peaked at time = 6 s after picture onset (see Fig. 2), which corresponds to the delay of the hemodynamic response. Analyses of variance of the responses during the corresponding time frame ($t=5$ –8 s) revealed that, at this point, all regions of interest (except for the locus coeruleus and parahippocampal cortex) responded more strongly to negative pictures than to neutral pictures (all $p<.05$). Initially, negative pictures thus engaged these brain regions to a greater degree than neutral pictures. The results of these analyses are given in Table 2.

Because the task was presented 4 s after picture onset, effects of task load on the response to the picture would be expected at about 10 s and onwards. In line with this, activation in cognitive processing regions, such as the right dorsolateral frontal cortex, peaked at time = 10 s, e.g. during the third time frame ($t=9$ –12 s; see Fig. 2). Moreover, if task load indeed modulates the temporal unfolding of emotional responses, this may imply an interaction effect between picture valence and task load on participants' brain activity during this time frame. We indeed found statistically significant interactions between picture valence and task load in several regions during the third time frame ($t=9$ –12 s) in right dorsolateral frontal cortex ($F(1, 16)=4.69$, $p=.046$), right superior parietal cortex ($F(1, 16)=5.28$, $p=.035$), the left dorsal occipital cortex ($F(1, 16)=5.16$, $p=.037$), the left amygdala ($F(1, 16)=4.74$, $p=.045$), the right amygdala ($F(1, 16)=4.82$, $p=.043$), and the right inferior insula ($F(1, 16)=7.70$, $p=.014$) (see Table 3 for an overview).

The effect of task load on brain responses to negative pictures

Recall that we predicted neural responses in emotion regions to negative pictures to be attenuated when followed by a complex task (high load) rather than a simple task (low load), while we expected neural responses in cognitive processing regions to display the opposite pattern. To test these predictions, we directly compared the responses to negative pictures followed by either a complex or a simple arithmetic task in the regions that showed interactions between picture valence and task complexity in the remaining two time frames following task onset (i.e. the third and the fourth time frame). Note that these comparisons result in a positive t -value if the response is greater in the high load than the low load trials, and a negative t -value if the response is greater in the low load trials.

During the third time frame ($t=9$ –12 s), the left amygdala ($t(16)=-2.91$, $p=.010$) and the right amygdala ($t(16)=-1.97$, $p=.066$) showed weaker responses to negative pictures followed by a complex task rather than a simple task. During the fourth time frame ($t=13$ –16 s), the left amygdala ($t(16)=-4.29$, $p=.001$), right amygdala ($t(16)=-5.11$, $p=.0001$), and right insula ($t(16)=-2.25$, $p=.039$) showed weaker responses to negative pictures followed by a complex task rather than a simple task (see Table 4). These regions are known to be involved in the processing of emotional stimuli (Ochsner et al., 2004; Phan et al., 2002). Thus, increased cognitive load led to reduced activation in emotion circuits.

During the third time frame ($t=9$ –12 s), right dorsolateral frontal cortex ($t(16)=7.94$, $p<.0001$), right superior parietal cortex ($t(16)=6.62$, $p<.0001$), and left dorsal occipital cortex ($t(16)=5.01$, $p=.0001$) showed greater activation to negative pictures followed by a complex task rather than a simple task. During the fourth time frame ($t=13$ –16 s), right dorsolateral frontal cortex ($t(16)=3.43$, $p=.003$), right superior parietal cortex ($t(16)=2.43$, $p=.027$), and left dorsal occipital cortex ($t(16)=2.52$, $p=.023$) showed greater activation to negative

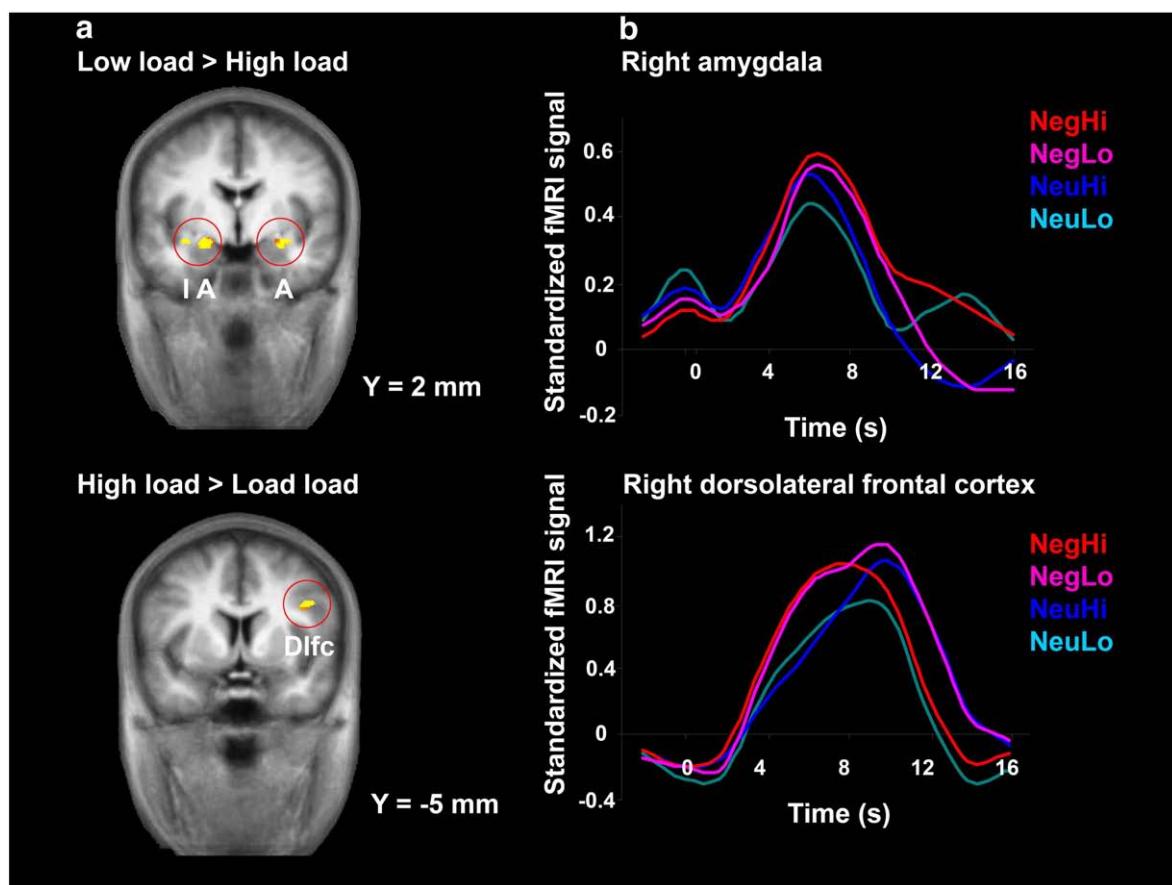


Fig. 2. (a) Two coronal views of the averaged brain ($n=17$) in Talairach space on which areas are displayed that responded significantly to negative pictures ($p<.05$, corrected for multiple comparisons) and in which activity was either greater when task load was low compared to high (upper row; amygdalae [A] and right inferior insula [IA]), or in which activity was greater when task load was high compared to low (lower row; right dorsolateral frontal cortex [Dlfc]). (b) Deconvolved averaged timecourses of the brain responses in the right amygdala and right dorsolateral frontal cortex for each trial type; negative (Neg) or neutral (Neu) pictures that were followed by either a high (Hi) or a low (Lo) load task. Picture onset was at time = 0 s, task onset was at time = 4 s. Due to the hemodynamic delay, effects of pictures valence are observed from time = 6 s onwards and effects of task load from time = 10 s onwards.

pictures followed by a complex task rather than a simple task (see Table 4). These regions are known to be involved in cognitive processing and to respond to increasing task demands (De Fockert et al., 2001; Duncan and Owen, 2000; Prabhakaran et al., 2000; Rypma et

al., 1999). Thus, as expected, performing complex arithmetic sums engaged brain regions that support cognitive processes.

Recall that we found an effect of task complexity on participants' response times. To investigate whether the effects of task load could be attributed to response time differences between the complex and the simple arithmetic equations, we repeated the analyses with participant's response time difference between complex and simple trials as a covariate. These analyses did not reveal any differential findings, except for the left amygdala during the third time frame ($t=9-12$ s). When entered as a covariate, response time differences partially accounted for the effect of task complexity on left amygdala responses, $F(1,15)=7.22$, $p=.017$. However, a 2 (task complexity)

Table 2

Regions of interest showing a greater response to negative than to neutral pictures between 5 and 8 s after picture onset ($p<.05$)

Brain region (Brodmann area, hemisphere)	Talairach coordinates (mm)			Volume (ml)	F_{val}
	x	y	z		
Dorsolateral frontal cortex (BA 6/44, right)	40	2	30	0.210	61.56*
Superior parietal cortex (BA 7, right)	24	-59	53	0.331	13.63*
Fasciculus uncinatus (BA 34, left)	-28	7	-11	0.353	43.23*
Dorsal occipital cortex (left)	-22	-79	19	13.979	12.13*
Dorsal occipital cortex (right)	23	-77	18	13.129	8.63
Ventral occipital cortex (left)	-22	-64	-5	30.286	6.47
Medial anterior temporal cortex (BA 38, right)	34	-4	-23	0.221	8.53
Amygdala (left)	-19	-5	-9	0.211	51.89*
Amygdala (right)	22	-8	-10	0.450	5.32
Inferior insula (left)	-31	-1	-9	465	10.94*
Inferior insula (right)	31	-3	-9	0.173	5.32
Pulvinar (left)	-18	-24	0	0.118	25.89*
Pulvinar (right)	19	-24	0	0.986	8.31
Superior colliculus (left)	-6	-24	0	0.934	77.66*
Superior colliculus (right)	6	-24	0	0.800	57.21*

Note. $F_{\text{val}}=F$ -value of picture valence, all negative>neutral. Degrees of freedom are 1 and 16, * $p<.05$.

Table 3

Regions of interest showing a picture valence \times task load interaction between 9 and 12 s after picture onset

Brain region (Brodmann area, hemisphere)	Talairach coordinates (mm)			Volume (ml)	F_{int}
	x	y	z		
Dorsolateral frontal cortex (BA 6/44, right)	40	2	30	0.210	4.69
Superior parietal cortex (BA 7, right)	24	-59	53	0.331	5.28
Dorsal occipital cortex (left)	-22	-79	19	13.979	5.16
Amygdala (left)	-19	-5	-9	0.211	4.74
Amygdala (right)	22	-8	-10	0.450	4.82
Inferior insula (right)	31	-3	-9	0.173	7.70

Note. $F_{\text{int}}=F$ -value of picture valence \times task load interaction. Degrees of freedom are 1 and 16, * $p<.05$.

Table 4

Regions of interest displaying an effect of task load on responses to negative pictures during the third and fourth time frame ($t=9\text{--}12$ s, and $t=13\text{--}16$ s, respectively)

Brain region (Brodmann area, hemisphere)	Talairach coordinates (mm)			Volume (ml)	T_{neg}	
	x	y	z			
High load>low load						
Dorsolateral frontal cortex (BA 6/44, right)	40	2	30	0.210	7.94**	3.43**
Superior parietal cortex (BA 7, right)	24	-59	53	0.331	6.62**	2.43*
Dorsal occipital cortex (left)	-22	-79	19	13.979	5.01**	2.52*
Low load>high load						
Amygdala (left)	-19	-5	-9	0.211	-2.91*	-4.29**
Amygdala (right)	22	-8	-10	0.450	-1.97†	-5.11**
Inferior insula (right)	31	-3	-9	0.173	-1.11	-2.25*

Note. T_{neg} =t-value of effect of task load after negative pictures; with positive t meaning greater response to high than to low load, and negative t meaning greater response to low than to high load. Degrees of freedom=16, ** $p<.01$, * $p<.05$, † $p<.1$.

ANCOVA analysis of the brain responses in the negative trials still revealed a significant effect for task complexity ($F(1,16)=5.01$, $p=.040$). Accordingly, the effects of task complexity could not be fully accounted for by response time differences but instead may be the result of the differential cognitive involvement in the task.

Interrelation of brain regions involved in emotion and task processing

We examined the interrelations of neural activity in emotion regions (i.e. the amygdala and right insula) on the one hand and regions implicated in the arithmetic task (i.e. right dorsolateral frontal cortex, right superior parietal cortex, left dorsal occipital cortex) on the other hand. To this end, we first computed an index of the amount of modulation of picture valence by task load. We did this by calculating the difference between responses to negative pictures followed by a simple task and neutral pictures followed by a simple task, minus the difference between responses to negative pictures followed by a complex task and neutral pictures followed by a complex task, separately for each participant, ROI, and fMRI time frame. Note that because of the double subtraction, the sign of the task load effect is lost. We then correlated these indices of the modulation of the valence effect across participants between ROIs, separately for each time frame.

We found that activation in left amygdala correlated with activation in right dorsolateral frontal cortex ($r=.52$, $p=.027$) and left dorsal occipital cortex ($r=.47$, $p=.049$) during the third time frame ($t=9\text{--}12$ s) following picture onset. Moreover, activation in right inferior insula correlated with activation in right dorsolateral frontal cortex during the third (displaying a trend; $r=.43$, $p=.075$) and fourth time frame following picture onset ($r=.47$, $p=.049$). Whereas activation in right dorsolateral frontal cortex and left dorsal occipital cortex was greater in response to high compared to low task load, activation in limbic regions was smaller in response to high compared to low task load (see previous section, Table 4 and Fig. 2). In line with predictions, there thus seemed to be a systematic relationship between activity in regions implicated in emotional processing (amygdala, right insula) and activity in regions implicated in cognitive processing (right dorsolateral frontal cortex, left dorsal occipital cortex) which is consistent with the idea that these systems operate in a coordinated manner (Pessoa, 2008). However, given that we could not control for collinearity and that the reported correlations were only moderately strong, we suggest these coactivation findings be interpreted with caution.

Gender differences

Research has reported gender differences in brain responses to standardized emotional stimuli (Canli et al., 2002; Mackiewicz et al.,

2006; Wrase et al., 2003). To investigate any effects of gender, we repeated all above analyses for the responses of the female participants only ($n=13$). Excluding the male participants from the analyses did not alter the observed response patterns.

Discussion

The present research examined how cognitive load modulates the unfolding of the emotional brain response. Importantly, in the present paradigm, the emotional stimulus and the cognitive load were presented in succession. Thus, participants solved an arithmetic equation following each neutral or negative picture. This temporal separation made it possible to investigate the dynamic unfolding of the interplay between cognitive and emotional neural structures. Moreover, by inducing cognitive load following rather than during picture display, we could examine the effects of cognitive load on the unfolding of the emotional brain response, while controlling for differences in visual processing strategies (Van Reekum et al., 2007). Previous research has proposed that task load may 'short-circuit' emotional processing in the brain, for example by deploying visual attention, such that activity in emotional brain regions is inhibited altogether. In the present research, however, we find that even when emotional circuits have already been engaged, performing a demanding task may still attenuate processing in the emotional brain.

The present findings indicate that both regions involved during the arithmetic task (right dorsolateral frontal cortex, right superior parietal cortex and left dorsal occipital cortex) and emotion regions (bilateral amygdala, right insula) initially showed greater activity in response to negative pictures than in response to neutral pictures. Consistent with previous work (Cohen et al., 1997; Prabhakaran et al., 2000), high task load further resulted in an increase in activity in the regions implicated in the arithmetic task (right dorsolateral frontal cortex, right superior parietal cortex, dorsal occipital cortex). More importantly, task load also resulted in a decrease in brain regions involved in emotion processing (bilateral amygdala, right insula). Finally, during task performance, activity in right dorsolateral frontal cortex, and, to a lesser extent, left dorsal occipital cortex, was related to activity in emotion regions (left amygdala, right insula). Together, these findings suggest that emotional and cognitive circuits in the brain operate in a coordinated manner to deal with changing task demands.

The present findings go beyond a simple reciprocal modulation of cognitive and emotion circuits in the brain, in that increases in activity in cognitive brain regions not necessarily resulted in decreases in activity in emotional brain regions. That is, during picture display, activity in right dorsolateral frontal cortex, right superior parietal cortex, and left dorsal occipital cortex, as well as limbic regions, i.e. the bilateral amygdala and the insulae, was greater in response to negative pictures than to neutral pictures. Only when participants performed the arithmetic equations neural responses in these regions began to differentiate. Whereas activity in cognitive regions increased even more in right dorsolateral frontal cortex, right superior parietal cortex, and left dorsal occipital cortex, performing the arithmetic equations resulted in a decrease in neural activity in limbic regions.

Coactivation analysis revealed that these opposite response patterns in higher cortical versus limbic regions were related, suggesting that the more right dorsolateral frontal cortex and left dorsal occipital cortex were engaged by the arithmetic task, the more emotional brain responses were attenuated. This is in line with previous research reporting a suppression of activity in limbic regions by (frontal) cortical regions during higher cognitive processes (Drevets and Raichle, 1998). However, given the correlational nature of our analyses, the causality of these relationships cannot be determined. We therefore suggest the present findings be interpreted with caution.

In line with previous findings (Erk et al., 2007), right dorsolateral frontal cortex was "shared" by task-related and emotional processing, such that this region was particularly engaged when negatively

valenced pictures were followed by complex arithmetic tasks. Accordingly, our results reveal an integration effect for processing negative emotion and task load in right dorsolateral frontal cortex. One potential explanation for these findings is that the right dorsolateral frontal cortex is engaged more with increasing task load, in order to sustain priority to processing of the central task at the cost of the further processing of emotionally salient, but task irrelevant negative stimuli (see Erk et al., 2007, for a similar argument). Theorists have proposed that the lateral frontal/prefrontal cortex may function as a key neural substrate of the central bottleneck of information processing (Dux et al., 2006; Herath et al., 2001; Marois and Ivanoff, 2005). As such, the lateral frontal cortex may be critical for mental operations such as cognitive control, decision-making, and modality-independent selection of task-relevant information (Badre et al., 2005; Brass et al., 2005; Bunge et al., 2003). The present findings suggest that the role of the lateral frontal cortex in controlling information processing may extend beyond the cognitive domain, and may similarly control the processing of emotional information.

Cognitive emotion regulation strategies may build upon more general information processing systems such as working memory and cognitive control that typically engage (frontal) cortical regions such as the dorsolateral frontal cortex, the superior parietal cortex and the dorsal occipital cortex (Ochsner and Gross, 2008). Indeed, a recent study revealed a role of the lateral frontal cortex in inhibiting emotional distraction in healthy adults during a working memory task (Johnson et al., 2005; Dolcos and McCarthy, 2006). Performing a working memory task engaged the lateral frontal cortex to a greater degree when the participants were distracted by negative emotional pictures rather than neutral or scrambled pictures. Moreover, participants who displayed greater activity to emotional distractors in the lateral frontal cortex judged emotional distractors as less distracting and less emotional.

The present findings potentially have important implications for psychopathologies that are characterized by both emotional and cognitive deficits, like depression and anxiety (Harvey et al., 2005; Luu et al., 1998; Mayberg et al., 1999). Emotional stimuli can have a lasting effect on amygdala activity (Cuthbert et al., 2000), specifically in depressed individuals (Siegle et al., 2002). Moreover, research has shown that people are inclined to ponder over a negative experience, which may subsequently intensify and prolong people's negative emotional states (Nolen-Hoeksema et al., 1993). Placing people in a quiet environment in order to let them 'cool down' from an emotional response may therefore not necessarily result in a more neutral state of mind. Although more research is needed in this area, it is conceivable that cognitively demanding tasks may eventually be used as a therapeutic tool. Having people perform an engaging task may alleviate the intensity of acute emotional responses, such that people become more receptive to long-term therapeutic interventions.

More broadly speaking, the present work attests to the close coordination between cognitive and emotional functioning. Whereas both cognitive and emotional brain circuits are involved in processing negative pictures, our findings demonstrated that further engagement of cognitive circuits by a subsequent task go hand in hand with a decrease in activity in emotional circuits (Drevets and Raichle, 1998; Keightley et al., 2003; Mayberg et al., 1999), especially when processing load of the task is high. Depending on the threats and challenges that people face, cognitive and emotional systems are recruited in a flexible manner, such that people can deal effectively with the ever changing demands of their environment.

References

- Ashcraft, M.H., Kirk, E.P., 2001. The relationships among working memory, math anxiety, and performance. *J. Exp. Psychol. Gen.* 130, 224–237.
- Badre, D., Poldrack, R.A., Pare-Blairev, E.J., Insler, R.Z., Wagner, A.D., 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47, 907–918.
- Blair, K.S., et al., 2007. Modulation of emotion by cognition and cognition by emotion. *Neuroimage* 35, 430–440.
- Brass, M., Derrfuss, J., Forstmann, B., von Cramon, D.Y., 2005. The role of the inferior frontal junction area in cognitive control. *Trends Cogn. Sci.* 9, 314–316.
- Bunge, S.A., Kahn, I., Wallis, J.D., Miller, E.K., Wagner, A.D., 2003. Neural circuits subserving the retrieval and maintenance of abstract rules. *J. Neurophysiol.* 90, 3419–3428.
- Canli, T., Desmond, J.E., Zhao, Z., Gabrieli, J.D.E., 2002. Sex differences in the neural basis of emotional memories. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10789–10794.
- Cohen, J.D., Perlstein, W.M., Braver, T.S., Nystrom, L.E., Noll, D.C., Jonides, J., et al., 1997. Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608.
- Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J., 2000. Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol. Psychol.* 52, 95–111.
- De Fockert, J.W., Rees, G., Frith, C.D., Lavie, N., 2001. The role of working memory in visual selective attention. *Science* 291, 1803–1806.
- Dolcos, F., McCarthy, G., 2006. Brain systems mediating cognitive interference by emotional distraction. *J. Neurosci.* 26, 2072–2079.
- Drevets, W.C., Raichle, M.E., 1998. Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: implications for interactions between emotion and cognition. *Cogn. Emot.* 12, 353–385.
- Duncan, J., Owen, A.M., 2000. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483.
- Dux, P.E., Ivanoff, J.G., Asplund, C.L., Marois, R., 2006. Isolation of a central bottleneck of information processing with time-resolved fMRI. *Neuron* 52, 1109–1120.
- Erber, R., Tesser, A., 1992. Task effort and the regulation of mood – the absorption hypothesis. *J. Exp. Soc. Psychol.* 28, 339–359.
- Erthal, F.S., De Oliviera, L., Mocaiber, I., Pereira, M.G., Machado-Pinheiro, W., Volchan, E., 2005. Load-dependent modulation of affective picture processing. *Cogn. Affect. Behav. Neurosci.* 5, 388–395.
- Erk, S., Kleczar, A., Walter, H., 2007. Valence specific regulation effects in a working memory task with emotional context. *Neuroimage* 37, 623–632.
- Friston, K.J., Holmes, A.P., Price, C.J., Buchel, C., Worsley, K.J., 1999. Multisubject fMRI studies and conjunction analyses. *Neuroimage* 10, 385–396.
- Gerin, W., Davidson, K.W., Christenfeld, N.J.S., Goyal, T., Schwartz, J.E., 2006. The role of angry rumination and distraction in blood pressure recovery from emotional arousal. *Psychosom. Med.* 68, 64–72.
- Glynn, L.M., Christenfeld, N., Gerin, W., 2002. The role of rumination in recovery from reactivity: cardiovascular consequences of emotional states. *Psychosom. Med.* 64, 714–726.
- Goebel, R., Muckli, L., Zanella, F.E., Singer, W., Stoerig, P., 2001. Sustained extrastriate cortical activation without visual awareness revealed by fMRI studies of hemianopic patients. *Vis. Res.* 41, 1459–1474.
- Gross, J.J., 2001. Emotion regulation in adulthood: timing is everything. *Curr. Dir. Psychol. Sci.* 10, 214–219.
- Harvey, P.O., Fossati, P., Pochon, J.B., Levy, R., LeBastard, G., Leherier, S., et al., 2005. Cognitive control and brain resources in major depression: an fMRI study using the n-back task. *Neuroimage* 26, 860–869.
- Herath, P., Klingberg, T., Young, J., Amunts, K., Roland, P., 2001. Neural correlates of dual task interference can be dissociated from those of divided attention: an fMRI study. *Cereb. Cortex* 11, 796–805.
- Johnson, M.K., Raye, C.L., Mitchell, K.J., Greene, E.J., Cunningham, W.A., Sanislow, C.A., 2005. Using fMRI to investigate a component process of reflection: prefrontal correlates of refreshing a just-activated representation. *Cogn. Affect. Behav. Neurosci.* 5, 339–361.
- Keightley, M.L., Winocur, G., Graham, S.J., Mayberg, H.S., Hevenor, S.J., Grady, C.L., 2003. An fMRI study investigating cognitive modulation of brain regions associated with emotional processing of visual stimuli. *Neuropsychologia* 41, 585–596.
- Lang, P.J., Bradley, M.M., Cuthbert, B.N., 2001. International Affective Picture System (IAPS): Instruction Manual and Affective Ratings. Technical report A-5. The Center for Research in Psychophysiology, University of Florida.
- Lyubomirsky, S., Tucker, K.L., Caldwell, N.D., Berg, K., 1999. Why ruminators are poor problem solvers: clues from the phenomenology of dysphoric rumination. *J. Pers. Soc. Psychol.* 77, 041–1060.
- Luu, P., Tucker, D.M., Berryberry, D., 1998. Anxiety and the motivational basis of working memory. *Cogn. Ther. Res.* 22, 577–594.
- Mackiewicz, K.L., Sarinopoulos, I., Cleven, K.L., Nitschke, J.B., 2006. The effect of anticipation and the specificity of sex differences for amygdala and hippocampus function in emotional memory. *Proceedings of the National Academy of Sciences of the United States of America* 103, 14200–14205.
- Marois, R., Ivanoff, J., 2005. Capacity limits of information processing in the brain. *Trends Cogn. Sci.* 9, 296–304.
- Mayberg, H.S., Liotti, M., Brannan, S.K., McGinnis, S., Mahurin, R.K., Jerabek, P.A., et al., 1999. Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am. J. Psychiatry* 156, 675–682.
- Mitchell, D.G., Nakic, M., Fridberg, D., Kamel, N., Pine, D.S., Blair, R.J., 2007. The impact of processing load on emotion. *Neuroimage* 34, 1299–1309.
- Morrow, J., Nolen-Hoeksema, S., 1990. Effects of responses to depression on the remediation of depressive affect. *J. Pers. Soc. Psychol.* 58, 519–527.
- Nolen-Hoeksema, S., Morrow, J., Fredrickson, B.L., 1993. Response styles and the duration of episodes of depressed mood. *J. Abnorm. Psychology* 102, 20–28.
- Ochsner, K.N., Ray, R.D., Cooper, J.C., Robertson, E.R., Chopra, S., Gabrieli, J.D.E., et al., 2004. For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. *Neuroimage* 23, 483–499.

- Ochsner, K.N., Gross, J.J., 2005. The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249.
- Ochsner, K.N., Gross, J.J., 2008. Cognitive emotion regulation: insights from social cognitive and affective neuroscience. *Curr. Dir. Psychol. Sci.* 17, 153–158.
- Öhman, A., 2007. Has evolution primed humans to “beware the beast”? *Proc. Natl. Acad. Sci. U. S. A.* 104, 16396–16397.
- Okon-Singer, H., Tzelgov, J., Henik, A., 2007. Distinguishing between automaticity and attention in the processing of emotionally significant stimuli. *Emotion* 7, 147–157.
- Pessoa, L., 2008. On the relationship between emotion and cognition. *Nat. Rev. Neurosci.* 9, 148–158.
- Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11458–11463.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage* 16, 331–348.
- Prabhakaran, V., Narayanan, K., Zhao, Z., Gabrieli, J.D.E., 2000. Integration of diverse information in working memory within the frontal lobe. *Nat. Neurosci.* 3, 85–90.
- Rypma, B., Prabhakaran, V., Desmond, J.E., Glover, G.H., Gabrieli, J.D.E., 1999. Load-dependent roles of frontal brain regions in the maintenance of working memory. *Neuroimage* 9, 216–226.
- Sapolsky, R.M., 1994. Why Zebras Don't Get Ulcers: A guide to Stress, Stress-related Diseases and Coping. WH Freeman, New York.
- Siegle, G.J., Steinhauer, S.R., Thase, M.E., Stenger, V.A., Carter, C.S., 2002. Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biol. Psychiatry* 51, 693–707.
- Talairach, J., Tournoux, P., 1988. Co-planar Stereotaxic Atlas of the Human Brain. M. Rayport, Transl, Thieme, New York.
- Van Dillen, L.F., Koole, S.L., 2007. Clearing the mind: a working memory model of distraction from negative feelings. *Emotion* 7, 715–723.
- Van Dillen, L.F., and Koole, S.L., in press. How automatic is “automatic vigilance”? The role of working memory in interference of negative information. *Cognition and Emotion*.
- Van Reekum, C.M., Johnstone, T., Urry, H.L., Thurow, M.E., Schaefer, H.S., Alexander, A.L., Davidson, R.J., 2007. Gaze fixation predicts brain activation during the voluntary regulation of picture-induced negative affect. *Neuroimage* 36, 1041–1055.
- Wräse, J., Klein, S., Grüsser, S.M., Hermann, D., Flor, H., Mann, K., et al., 2003. Gender differences in the processing of standardized emotional visual stimuli in humans: a functional magnetic resonance imaging study. *Neurosci. Lett.* 348, 41–45.

A Neurocognitive Model of Advertisement Content and Brand Name Recall

Antonio G. Chessa

Statistics Netherlands (CBS), P.O. Box 4000, 2270 JM, Voorburg, The Netherlands, antoniogchessa@yahoo.com

Jaap M. J. Murre

Department of Psychonomics, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands,
jaap@murre.com

We introduce a new (point process) model of learning and forgetting, inspired by the structures of the brain, that we apply to model long-term memory for advertising and brand name recall. Recall-probability functions derived from the model are tested with classic data by Zielske (1959), as well as advertisement content and brand name recall data of a Dutch study that tracked over 40 campaigns of TV commercials. Data fits and cross-validation results indicate that the recall functions serve as a good first approximation for aggregate behavior. The shapes of optimal GRP schedules, which are obtained by maximizing a recall measure, are strongly related to the model parameters and corresponding memory processes. Comparisons with existing models in the literature indicate that a neurobiologically motivated model may give a more realistic description of memory for advertisements.

Key words: advertising; memory; impact; scheduling; bursting; dripping; massed and spaced learning; point processes

History: This paper was received May 28, 2003, and was with the authors 27 months for 3 revisions processed by Joel Huber.

Introduction

Should advertisements be concentrated in time, spread out evenly, or distributed according to a mixture of these two strategies (e.g., Merzereau and Battais 2000)? A similar question is studied in the field of memory psychology: How should we learn new material such as foreign vocabulary words, according to a *massed* (bursting) or *spaced* (dripping) schedule? Most studies conclude that spaced learning is superior; others report that the optimal performance is obtained by some combination of massed and spaced learning (e.g., Baddeley and Longman 1978, Bahrick et al. 1993, Glenberg 1976, Glenberg and Lehmann 1980, Rumelhart 1967).

Memory performance for advertising material shows results similar to those obtained in memory psychology. Burtt and Dobell (1925) were the first to study the recall of advertisements; their results indicated a form of dependence between the interpretation lag and retention lag. The classic study by Zielske (1959) ranged over a period of one year, during which the recall of printed advertisements was measured during 13 repeated mailings that were sent to subjects either every week or every four weeks. The spaced, four-week schedule proved to produce a significantly higher memory performance. Although the superiority of the spaced schedule is evident from Zielske's study, these findings are not sufficient to assist decision makers in selecting optimal advertising

schedules. Is every four weeks the optimal schedule, or is the optimum located between one and four weeks, or at a larger spacing? Is this universally true, or does it depend on the brand and campaign?

The similarity between the results on this so-called spacing effect in the psychological and advertising literature, and the difficulty with extrapolating these empirical results to real-world situations, has motivated us to develop a mathematical memory model, the *memory chain model*. This model can be seen as an abstraction and extension of a neurological model of long-term memory by our group (Murre 1996; Murre et al. 2001; Meeter and Murre 2004, 2005; Murre et al. 2001; Talamini et al. 2005). In this paper, we will use the model to derive expressions for advertisement content and brand name recall. These recall functions will be fitted and validated on about 80 data sets. After the validation, we will use the fitted recall measures to optimize advertisement scheduling. Finally, we will study to what extent schedule type and short-term and long-term memory parameters correlate. We emphasize here that we will focus only on improving advertisement scheduling, not on improving the advertisement message.

In the next section, we will describe the most relevant details of the memory chain model. Our approach has a number of novel aspects: (1) the model formalizes neurobiological processes based on recent findings, and (2) we use a single model of memory to

derive different expressions for learning and forgetting of both advertisement content and brand name recall. The memory chain model has already been applied as such to hundreds of memory retention data sets (e.g., Chessà and Murre 2004, Janssen et al. 2005).

There are a few models in the psychological literature that have some similarity to the memory chain model (Atkinson and Shiffrin 1968, Murdock 1974). However, these models only describe situations with a single learning trial and subsequent forgetting, in which the short-term to long-term memory processes are rather restrictive. While the existing models are limited to one or two memory stores, the memory chain model can handle any number of stores. Two other types of modeling approaches are based on: (1) Markov chains (Murdock 1974), where the states represent the level of mastery of some item memorized; and (2) regression procedures for fitting functions to retention data (Rubin and Wenzel 1996). In the latter approach, the functions used are not derived from an underlying formal theory of memory, as is the case with the memory chain model, but are simply selected and fitted. Markov models only formalize the effect of learning at successive trials, but do not model forgetting between intermediate trials, which the memory chain model does take into account.

The data on which the recall functions for advertisement content and brand name will be tested and validated are presented in the third and fourth sections. The results of the data fits and optimization will be discussed in the discussion section, where we will

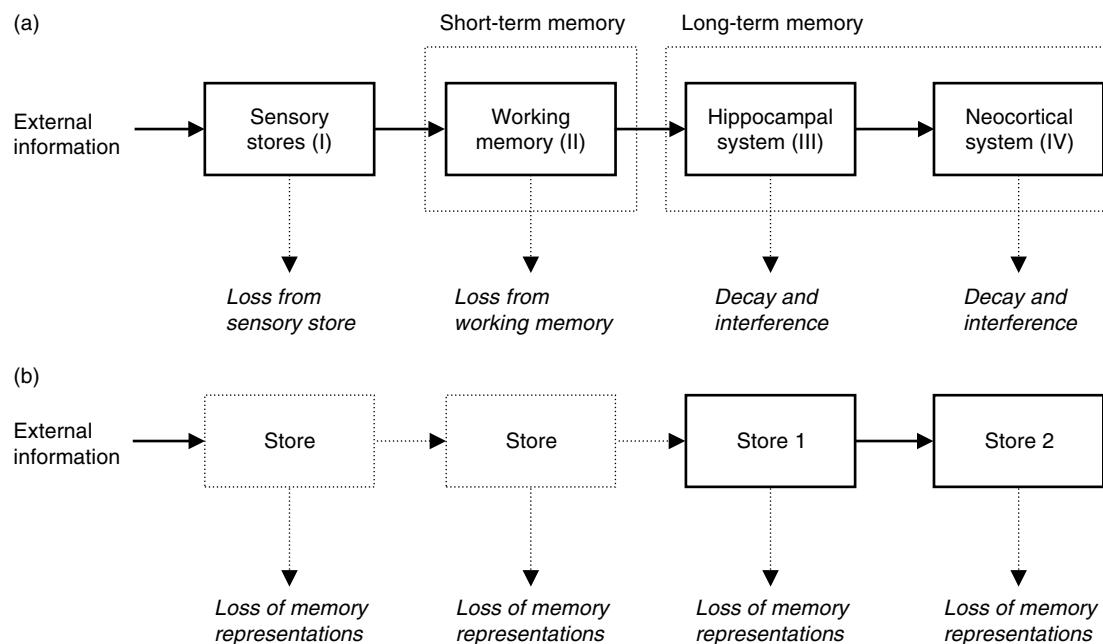
also compare the results with a model presented in Naik et al. (1998).

The Memory Chain Model

The memory model is built upon two main concepts. During exposure (e.g., to an advertisement), a memory is encoded as a number of *representations*, each of which captures characteristics of the memorized item. Representations may be activated over time by a memory cue, for instance, by a product category when a brand name has to be recalled. We model the process of activated memory representations in time as a point process (Daley and Vere-Jones 1988, Diggle 1983, Stoyan et al. 1987). We will consider a *Poisson point process*; that is, numbers of representations activated in time intervals have a Poisson distribution, and numbers in disjoint intervals are independent. Point processes have been used for modeling neural spike trains (Abeles 1991).

The second concept is motivated by neurobiological evidence that memories may be stored in one or more parts of the brain, which we will denote in abstract terms by *memory stores*. Two processes characterize each store: a process of memory decline or loss, and an induction process, which transfers representations from one store to another in a feed-forward fashion (Figure 1). A memory can be strengthened in a second store, for example, through induction by activation. This is an induction process where, for example, a rapidly decaying neural group activates neurons in a less rapidly decaying part of the brain. An exam-

Figure 1 (a) Storage Systems for a Memory at Different Time Scales, with Feed-Forward Induction Between and Decline Within Stores. (b) Abstract Representation Used in the Memory Chain Model



ple of the latter is the dorsolateral prefrontal cortex, which is assumed to hold a working memory representation for a few seconds to minutes (Goldman-Rakic 1992, 1995). From a neurobiological point of view, these processes are part of a cascade of induction and decline processes that take place at different time scales (McGaugh 2000).

We will use models with two stores, at most, in the fits to advertising data in order to limit the number of parameters. Because all the data fall in time ranges in the order of weeks and months, the two stores of the model will most likely reflect neural processes in the hippocampus (Store 1) and the neocortex (Store 2). There are no data at time lags shorter than a week, so that the short-term stores in the upper panel of Figure 1 are in fact covered by Store 1 in our model fits. It should therefore be noticed that there is not necessarily a one-to-one correspondence between the neural stores (Figure 1(a)) and the stores of the memory chain model (Figure 1(b)).

A Single Learning Trial

The memory chain model formalizes memory and retention as a process that consists of four stages, which are described below.

Encoding. During stimulus exposure, the number of memory representations eventually reaches a mean value, or *intensity*, μ_1 . Encoded memory representations are stored in the first memory store (i.e., Store 1). Representations are not labeled, so the model formalizes how much is stored. Memory representations could be seen as critical features or relevant details of an advertisement.

Storage. After encoding, a decline process is activated, for example, by spontaneous decay, by the overwriting of other learned items, by neural noise, or by other factors. An encoded memory representation is still available at time t since the end of a learning trial with probability denoted by $\tilde{r}_1(t)$, which will be called the *decline function* of Store 1. Memory in Store 1 is assumed to be a *nonhomogeneous Poisson process* with *intensity function* $r_1(t) = \mu_1 \tilde{r}_1(t)$. We will assume an exponential decline function, so that

$$r_1(t) = \mu_1 \exp(-a_1 t), \quad (1)$$

where $a_1 > 0$. Data obtained from laboratory experiments that intend to measure short-term memory decline through the classical Brown-Peterson learning and distraction task support an exponential decline (Peterson and Peterson 1959, Murdock 1961). Evidence at the neural level is offered by the long-term potentiation or synaptic growth and decline data in the hippocampus of young and old rats (Barnes and McNaughton 1980).

Memory representations may be induced from Store 1 to Store 2, where they form a new point process. This point process arises from a time-series of

inductions (which in some experiments can be thought of as rehearsals), of which the intensity is proportional to the intensity in the first store, that is, $\mu_2 \mu_1 \tilde{r}_1(t)$, where μ_2 is a (time-independent) induction rate. An induction that occurs at time τ induces a memory representation at time $t \geq \tau$ according to probability density function $a_2 \exp(-a_2(t - \tau))$ in Store 2, which we write as f . This function specifies the decline rate in the second store, which we assume to be smaller than in the first store ($a_2 < a_1$). The intensity function r_2 of the second store is a convolution of $\mu_2 \mu_1 \tilde{r}_1(t)$ and f , which yields the expression:¹

$$r_2(t) = \frac{\mu_1 \tilde{\mu}_2}{a_1 - a_2} (\exp(-a_2 t) - \exp(-a_1 t)), \quad (2)$$

where $\tilde{\mu}_2 = \mu_2 a_2$, for which we will continue to write μ_2 .

Retrieval. Memory stores are searched for evidence at a recall test. Stores can be searched either simultaneously or sequentially; it can be proved that both strategies lead to identical expressions for recall probability. The effectiveness of the search process and of the retrieval cue(s) used for the search is expressed by a single parameter q , which can be interpreted as the probability that a memory representation will be cued. Cueing has the following effect on the total intensity $r_1 + r_2$ over the two stores:²

$$r(t) = (r_1(t) + r_2(t))q. \quad (3)$$

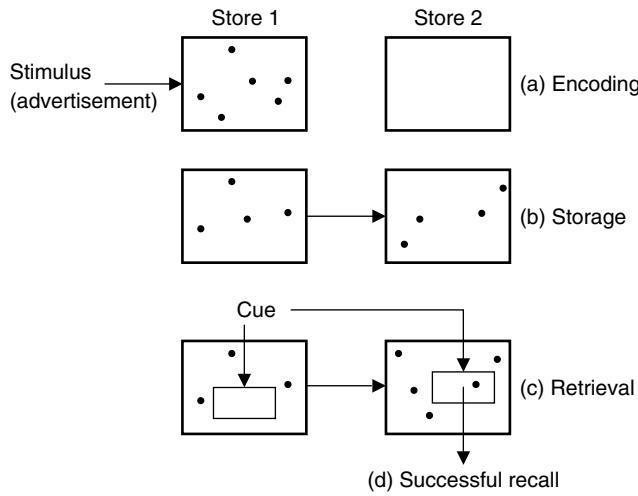
This model characteristic does not necessarily imply that cue effectiveness is constant over time. The retention functions that follow from our memory model are invariant under exponentially varying cue effectiveness. This assumption does not lead to additional degrees of freedom in the model, because cue effectiveness can be subsumed under the exponential memory decline functions of the stores. We will combine q with the encoding parameter μ_1 , for which we will continue to write μ_1 , because the effect of cueing is not varied systematically in the studies reported here.

Recall. Our model includes a fourth stage, which can be considered a decision process that relates retrieval to recall. The number of retrieved representations is compared with a threshold b . Unless stated otherwise, we assume that successful recall takes place if a store exists with at least one cued representation. In some cases, subjects require more than one representation for recall. From laboratory exper-

¹ For more details, see <http://www.neuromod.org/staff/murre/mcmforgettingdraft.doc>.

² The formalisms of memory decline and cueing, as shown in (1) and (3), are in fact *independent thinnings* of Poisson processes with intensities μ_1 and $r_1 + r_2$, respectively, which again result in Poisson processes with thinned intensities (Stoyan et al. 1987).

Figure 2 Illustration of Memory Processes in a Two-Store Model



Notes. (a) Memory representations (points) are generated in Store 1. (b) While memory declines in Store 1, representations are generated in a second store by induction. (c) A cue searches areas in both stores and finds a memory representation in Store 2. (d) A correct response will be given at a recall threshold of 1.

iments, we obtained a strong suggestion that a new stimulus has a higher threshold and thus requires more representations than a well-known stimulus does. From the properties of Poisson processes, it follows that the *recall-probability function* at retention lag t since stimulus exposure takes the form

$$p(t) = 1 - \sum_{n=0}^{b-1} \frac{(r(t))^n}{n!} \exp(-r(t)). \quad (4)$$

For the special case $b = 1$, we thus have $p(t) = 1 - \exp(-r(t))$. The concepts of memory encoding, storage, retrieval, and recall are illustrated in Figure 2.

Multiple Learning Trials

We assume that different learning trials give rise to independent Poisson processes, where memory representations generated by different learning trials follow the same storage and retrieval processes as in the single-trial case. The only opportunity for interaction of successive learning trials is at encoding. From a neurobiological point of view, it is rather uncommon that encoding increases linearly as learning continues. Rather, it will saturate, and therefore be limited by some upper bound. For example, the number of firing neural groups cannot become infinitely large.

We will, therefore, assume that the initial encoding may saturate at some finite value $r_{\max} \geq \mu_1$. Let v denote the learning rate per unit of learning time and l denote the duration of the first learning trial. We

then describe the initial encoding during the learning trial as follows:³

$$\mu_1(l) = r_{\max}(1 - e^{-vl/r_{\max}}). \quad (5)$$

When μ_1 is close to saturation, learning no longer has a significant effect, which is what we expect in memory psychology with prolonged massed learning.

The encoding intensity $\mu^{(L)}$ regarding the individual contribution of subsequent trials $L \geq 2$ is similar to (5). Let t_i , $i = 1, 2, \dots$, denote the presentation times of a series of learning trials. The encoding $\mu^{(L)}(l)$ of trial L with duration l is given by:

$$\mu^{(L)}(l) = \left(r_{\max} - \sum_{i < L} \mu^{(i)} \tilde{r}_1(t_L - t_i) \right) (1 - e^{-vl/r_{\max}}). \quad (6)$$

Notice that this expression is equal to $\mu_1(l)$, given by (5), multiplied by the relative intensity that can still be accommodated in Store 1 at trial L .

The recall probability $p(t)$ at a time lag t since the last trial in a series of L trials has the following form for recall threshold $b = 1$:

$$p(t) = 1 - \exp \left\{ - \sum_{i=1}^L \mu^{(i)} (\tilde{r}_1(t_L - t_i + t) + \tilde{r}_2(t_L - t_i + t)) \right\}, \quad (7)$$

where $\tilde{r}_2 = r_2/\mu_1$, with r_2 equal to (2). That is, \tilde{r}_2 denotes the decline function of Store 2. Other values of b lead to expressions similar to (4).

Existing Memories

It is easy to imagine situations where a memory has already been formed as a consequence of previous exposures to the same item, which may still exist during a new campaign. In our model, we account for such memory representations by assuming a memory that does not decline within the tracking period. We model this as a homogeneous Poisson process with (*base rate*) intensity μ_0 , which is assumed to be independent of the point processes induced by new campaigns. The implication of this additional process for recall probability is that the intensity μ_0 is summed with the intensity functions for new campaigns in the exponent of (7). A list of the parameters, functions, and variables in our model is given in Table 1.

We will now present two data sets to which we will apply our model. Recall-probability function (7) will be used in both applications as a first approximation to aggregate recall behavior concerning advertisement content and brand names. This means that we make

³ We use the term learning in a broad psychological sense, and it can be interchanged with encoding. Consumers could watch an advertisement passively (low v) or with a lot of attention (high v).

Table 1 Parameters, Functions, and Variables in the Memory Chain Model

	Theoretical meaning	Meaning/effect in applied context
Parameters		
μ_0	Intensity of homogeneous (base rate) Poisson process	Expected number of memory representations still left from previous campaigns
v	Learning rate per unit of learning time in time-continuous setting	Expected number of memory representations formed per GRP during encoding of ads
μ_1	Initial encoding (intensity after first trial)	Expected number of memory representations at the end of the first ad exposure or GRP pulse
r_{\max}	Upper bound on the intensity of Store 1 ($r_{\max} \geq \mu_1$)	Finite r_{\max} leads to decreasing encoding over ad exposures and to equal contributions otherwise
μ_2	Induction rate from Store 1 to Store 2	Constant transfer rate of memory representations of ads from Store 1 to Store 2
a_1, a_2	Decline rates in Stores 1 and 2, respectively	Relative loss of memory representations in Stores 1 and 2
b	Recall threshold	Minimum number of retrieved representations, e.g., critical features of ads, needed for successful recall
Functions		
$r(t)$	Intensity function summed over stores	Specifies expected number of memory representations taken over both stores as a function of test lag t
$r_1(t), r_2(t)$	Intensity functions of Stores 1 and 2	Same as $r(t)$, but for Stores 1 and 2 separately
$\tilde{r}_1(t), \tilde{r}_2(t)$	Decline functions of Stores 1 and 2	Express changes in the encoded number of memory representations for Stores 1 and 2 at test lags t
$\mu^{(L)}(l)$	Encoding for trial L with presentation time l	Expected number of memory representations for ad exposure L with l GRPs
$p(t)$	Recall-probability function	Probability of successfully recalling the content or the name of a brand in a commercial at test lag t
Variables		Meaning in applied context
t	Retention lag or interval	Time since last ad exposure
l	Learning or exposure time	GRPs in SPOT data
t_i	Time at which learning trial i is presented	Moment at which an ad is exposed for the i th time
L	Number of learning trials	Number of advertisements or GRP pulses in campaign

the simplifying assumption that memory processes are the same for different subjects. We make this choice because we propose a new model in this paper, and we want to find out to what extent this model is already able to fit the data with a limited number of parameters.

The Zielske Study

The Data

In the literature known to us, the only extensive study about the recall of printed advertising material for repeated advertisements is the classic study by Zielske (1959). Housewives were repeatedly sent the same printed advertisement, according to two schedules. One group received the advertisement once a week for 13 successive weeks, while the second group received the advertisement 13 times with a uniform spacing of four weeks. To obtain data about the recall of the advertisement in time, both groups were subdivided into smaller groups that were interviewed in specific weeks. The interviews were organized in such a way that data about learning and forgetting between and after the exposures were obtained. Simon (1979) analyzed the data and concluded that the one-week schedule reaches a higher peak in recall,

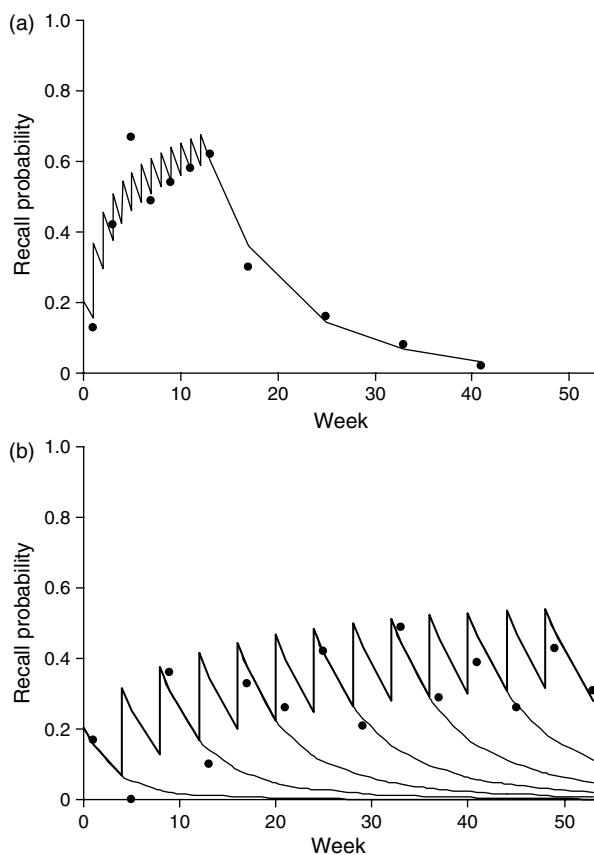
but that after about 17 weeks, recall improves for the four-week schedule and stays at a higher level after the end of the campaign. The average recall percentage calculated over a period of one year is higher for the four-week schedule. This implies that the spaced schedule would be preferred on two criteria, namely, (1) mid-term and long-term recall percentages after the end of an advertising campaign, and (2) the average recall percentage calculated over a period of one year.

The Fitted-Recall Functions

We fitted recall-probability function (7), which holds for recall threshold $b = 1$, and recall-probability functions with higher recall thresholds, which have the form of (4). A function was fitted simultaneously to both schedules. This means that we used the same parameter values in the fits to both schedules, because differences in recall for two different schedules should be explained by one underlying memory mechanism, and should merely represent the effect of intermediate forgetting between two successive advertisements.

The parameters were estimated by minimizing a chi-square statistic, which is the sum of conventional chi-square statistics computed for every advertising

Figure 3 Fits of a Two-Store Model with Saturated Learning to Zielske's Recall Data (Dots) of a Printed Advertisement



Notes. (a) Fit to the one-week advertising recall data: The first seven data points refer to 1, 3, 5, 7, 9, 11, and 13 advertisements received. The last four points denote forgetting after the 13th and final advertisement. (b) Fit to the four-week advertising recall data: The data points refer to 1, 3, 5, 7, 9, 11, and 13 advertisements, respectively, with each number of advertisements occurring twice in succession. The thin curves denote the course of forgetting after these numbers of advertisements.

week. The degrees of freedom are equal to the difference between the number of advertising weeks (i.e., data points) and the number of parameters estimated (Wickens 1982). From the minimum chi-square statistic we computed the size α for each fitted function at which a null hypothesis is not rejected. This approach was also carried out for the SPOT data in the next section.

Figure 3 shows the fit of a two-store model with saturated learning over advertisement exposures and a recall threshold $b = 2$. This was the only version of our model to fit the data for test sizes greater than 0.05 ($R^2 = 0.96$). The fits show the typical *saw-tooth* behavior, with recall peaks at each exposure and decline between successive exposures. Respondents were asked three different questions about the advertisement in order to test recall. This is reflected by the value of the recall threshold, which implies that they had to recall more than one feature of the advertisement.

The expected recall percentage over a one-year period can be calculated exactly for the two schedules simply by summing the expected recall over the weeks. On the basis of this criterion, the four-week schedule is one-and-a-half times more effective. An advantage of using a model is that other schedules can be defined and compared. We analyzed the expected yearly performance for schedules with both shorter and longer intervals, but the four-week schedule gave the highest one-year recall. Because of the saturation of advertisement memory, spaced advertising gives a better performance than massed advertising.

The SPOT Tracking Study

The Data

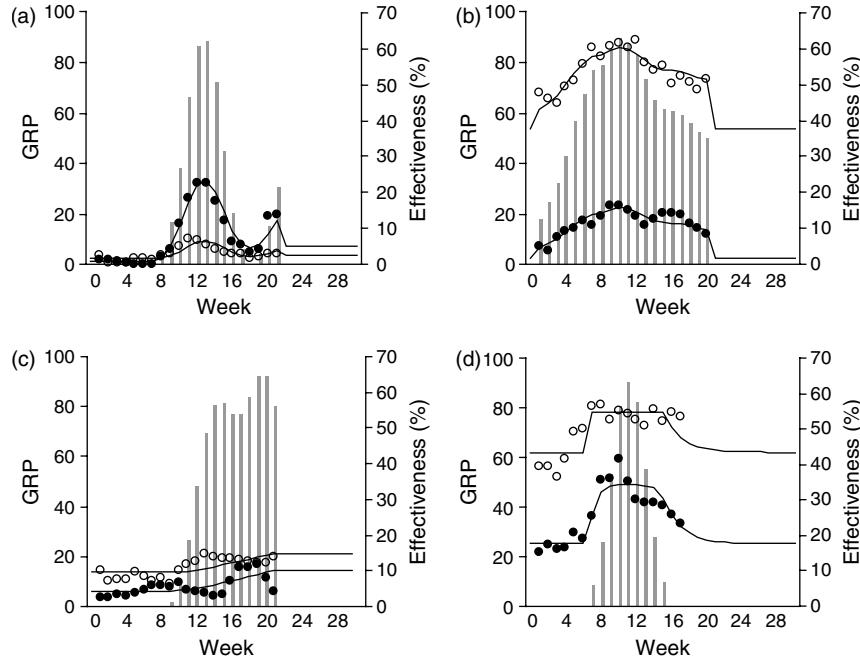
In 1997–1998, the Foundation for Promotion and Optimization of Television Advertising in the Netherlands (SPOT 1998) carried out an extensive study on the effectiveness of television advertising in the Netherlands. Product categories were considered with a large advertising budget, using television as the primary medium. A total of 42 brands within the categories "detergents" and "food" were tracked in campaigns that were scheduled in the period May 1997 until December 1997. (We do not mention brand names for reasons of confidentiality.) Data were collected through 50 phone calls per week for every brand. Only those respondents who did most of the shopping in their family and were older than 18 were interviewed. The results obtained were adjusted on the basis of age, education, residence, and type of household. Data about different advertising effectiveness measures were collected, of which we will analyze the following:

- *Impact:* This is proven recall about the contents of an advertisement. Respondents were asked, in an open-ended question, what they recalled about the TV commercial of a brand. The answers were coded according to a standard that was developed for this study and converted to a score between 0% and 100%;

- *Brand name recall:* This is the spontaneous recall of one or more brand names that belong to a specific product group. The interviewer named the product group, after which the respondent was asked to recall as many corresponding brand names as came to mind.

The data for these two effectiveness measures are presented in Figure 4 for a small selection of brands. Buying intention was measured as well in the SPOT study, but will be analyzed elsewhere, as this measure requires the integration of our model with a decision-analytic framework for comparing and ranking brands, which is beyond the scope of the present study.

Figure 4 Impact (Solid Circles) and Brand Name Recall (Open Circles) as a Function of GRPs (Gray Bars) for Four Brands (a: Meal, b: Detergent, c: Meal Sauce, d: Candy)



Notes. The model fits are shown as lines (thick line for impact, thin line for brand name recall): (a) a two-store model without saturation, rapid forgetting in Store 1, slow forgetting in Store 2, and a recall threshold equal to 1; (b) strong forgetting, no saturation; (c) no forgetting, no saturation; (d) strong saturation with slow forgetting.

The Fitted-Recall Functions

In Zielske's study, every individual was exposed to an advertisement once in an advertising week. In the SPOT study, individuals may have been exposed more than once, or never, to a commercial, which gives rise to a mean exposure frequency or GRP per week. We will approximate aggregate recall behavior by substituting learning time l in (5) and (6) by GRP for every advertising week or learning trial L .

Because existing commercials were also tracked, the base rate μ_0 for existing memories was included in the fits to both impact and brand name recall data. The remaining model assumptions with respect to impact are the same as for the model fitted to the Zielske data, that is, we fitted one-store and two-store models with the recall threshold as a free parameter. We fitted the same functions used for impact to brand name recall, subject to a number of restrictions. In contrast with the contents of the TV commercials, we assumed that all names of the 42 brands could be recalled easily, because they mostly consist of one word. The recall threshold b was therefore set equal to one in each of the 42 fits. The base rate μ_0 for existing memories may be different with respect to the base rate for impact, because a new commercial may be launched for a well-known brand, and because other media may also contribute to brand name recall. We also allowed the learning rate v to be different between brand names and other details of a commercial. The familiarity already gained with

a brand name through previous advertisements may decrease the contribution of TV commercials on its encoding. The remaining parameters were fixed at the same value obtained in the model fits to impact data. We thus have only two free parameters in the fits for brand name recall.

Fits to the SPOT data on impact and brand name recall are shown in Figure 4. On a total of 84 chi-square goodness-of-fit tests, three fits were rejected, at $\alpha = 0.05$, for both impact and brand name recall, which amounts to about 93% of nonrejections. We think that the rejections are caused by two factors: (1) nonuniform competition effects from other brands; (2) recall is not constant in the weeks before the TV-advertising campaign for certain brands, which indicates effects of possible advertising through other media. These effects are represented by the base rate parameter μ_0 , which we assumed to be constant in time.

A two-store model gave the best fit to the impact data for 11 brands. In Figure 4(a), the impact and brand name recall fits show that recall declines after the end of the campaign to values that are greater than the precampaign base rates, indicating memory consolidation to a second store. Of the remaining data sets, 28 were fitted with a one-store model, and three required only the stationary base rate process, which means that these three campaigns had no effect on impact.

A two-store model gave the best fit to brand name recall for seven campaigns, and 20 gave the best fit

with a one-store model. Fifteen brands only required the stationary base rate process. For these brands, the corresponding advertising campaigns did not have any effect on brand name recall. The results thus show numerous examples of TV commercials that improve impact but not brand name recall.

The question that we will try to answer in the next subsection is whether the memory processes that underlie the fitted-recall functions imply GRP schedules that improve the average impact and brand name recall over the tracking periods. For example, Figure 4(d) shows a brand with strong saturation, which is evidenced by the flat part of the impact and brand name recall functions during repeated advertising. Should one choose a schedule with an increased number of exposures, as is shown in Figure 4(d) during the first half of the campaign period; or a uniformly spaced schedule over a longer advertising period, given the decreasing contributions of repeated exposures to recall? Before answering such questions, we will subject the recall functions to a cross-validation study.

To assess the predictive performance of recall functions, we perform a cross-validation study, where a part of the data is used to estimate model parameters and the remaining data is predicted by using the functions fitted to the restricted data sets. We left out the data for the last five campaign weeks for every brand. Table 2 shows the values of three statistics for the recall functions fitted to the complete impact data sets and to the restricted data sets used for the cross validation. We also listed the results for the model proposed by Naik et al. (1998). Table 2 shows average values over all brands in each of four product classes.

We included R^2 because it is widely used in model fits. However, we would like to emphasize that it does not always give clear indications about the quality of a fit. A low R^2 value may correspond to test sizes $\alpha > 0.05$. This depends on factors such as the sample size per data point and the number of model parameters used in a fit. For example, a data set with small variation may be fitted well by a constant function, as additional parameters could decrease the test size α .

in a composite test. The result will then be a fit with a low R^2 value, which will increase with additional parameters, causing an opposite effect compared to α . Such fits also influenced the results in Table 2, as a comparison between the values for maximum α (which corresponds with the minimized chi-square statistic) and the number of rejected functions on the one hand, and the R^2 values on the other hand, illustrates.

The memory chain model gives better fits than the model by Naik et al. (1998). We will analyze and compare the two models in the discussion section. Table 2 also indicates that the model used for the cross validation gives results that are in close agreement with the model fitted to the complete data sets. The parameter estimates turned out to be stable when comparing the estimates for the restricted data and the full data. In more than 70% of the brands, the relative change in the parameter values remained within 10%. The fitted functions were hardly affected in the majority of the remaining cases, because the estimated parameter values were very small.

Optimization of Advertisement Scheduling

We used the recall-probability functions that were fitted to impact and brand name recall to investigate whether more effective campaigns exist with a different GRP scheduling, under the same total amount of GRP and campaign period. Different criteria can be chosen for optimization: impact and brand name recall in a specific target week (e.g., when advertising a movie premiere), or the average impact and brand name recall over a time interval (stimulate sales consistently). Here we focus on the latter.

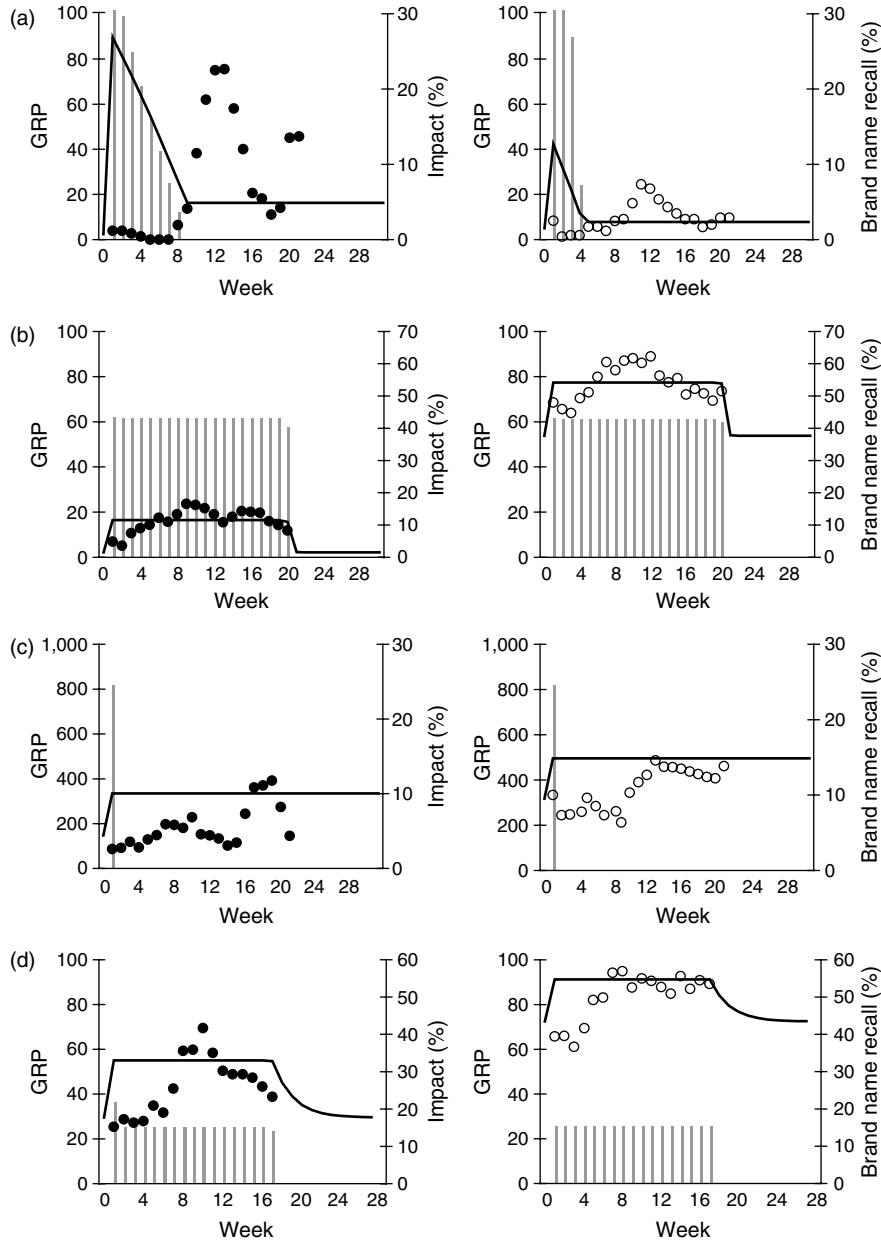
An increase of the average expected impact by more than 2% was found for six brands, and by more than 5% for two brands. Regarding brand name recall, two brands improve by 4%–5%, while the remaining brands improve up to 1%. A brand for which the average impact improves significantly under the optimal GRP distribution is a brand that is fitted with a one-store model with saturation and rapid forgetting (Figure 5(d)). For this brand, the average expected

Table 2 Values of Three Statistics for Three Models Fitted to the Impact Data (SPOT): Average Maximum Test Size α for Minimized Chi-Square Statistics, the Number of Rejected Functions ($\alpha = 0.05$), and Average R^2 (for Brands Grouped into Four Product Classes)

Product class	MCM (cross-validation)			MCM (complete fit)			NMS (complete fit)		
	Alpha	Rejected	R-square	Alpha	Rejected	R-square	Alpha	Rejected	R-square
Detergents (18)	0.56	1	0.44	0.55	0	0.50	0.46	4	0.49
Ice and coffee (8)	0.75	0	0.79	0.63	0	0.79	0.67	1	0.76
Meal products (11)	0.64	2	0.48	0.53	3	0.49	0.32	4	0.32
Candy (5)	0.59	0	0.67	0.59	0	0.82	0.50	0	0.81

Notes. The number of brands per product class is given between parentheses in the first column. MCM = memory chain model; NMS = model of Naik et al. (1998).

Figure 5 GRP Distributions (Gray Bars) for Maximized Average Expected Impact (Solid Line, Left) and Brand Name Recall (Solid Line, Right) over the Campaign Period



Notes. Data are for the same fitted functions as in Figure 4. The original impact and brand name recall data (dots) are also shown.

impact increases from 26.3% to 33%, while brand name recall increases from 50.1% to 54.8% under the optimal schedule. Instead of the bell-shaped GRP distribution shown in Figure 4(d), the memory processes that underlie this commercial and brand name rather suggest a uniformly spaced, dripping schedule. Significant increases were also found for the brands shown in Figures 5(a) and 5(c).

There is obviously a strong relation between the parameter values of a fitted-recall probability function and the type of optimal advertisement scheduling, which can be classified as follows. For brands that only consist of a stationary base rate process (i.e.,

no additional encoding during the campaign), dripping schedules with small amounts of GRP will be most cost effective. This situation arises three times for impact and 15 times for brand name recall. Brands that are memorized according to a one-store model with strong forgetting or that suffer from strongly saturated learning benefit most from a spaced distribution of GRP (dripping). Brands that are learned without saturation and that show very slow forgetting afterwards benefit most from a single burst at the beginning of a campaign (bursting). The same holds when the recall threshold increases (new brands or campaigns). In this way, the increased impact will

Table 3 Implications of Different Model Versions for Scheduling

Base rate	Learning	Forgetting	Recall threshold	Schedule
Exists	Very slow	Rapid	Low	Drip
			Low	Drip
	Saturated		Low	Drip
	Not saturated	Very slow		Burst
		Moderate, no consolidation	Low	Burst, then drip
	Not saturated	Rapid, with consolidation	Low High	Burst, then drip Burst

fall within the campaign period as much as possible. Other situations give rise to a mixture of bursting and dripping. Examples are brands and commercials that are memorized according to a two-store model without saturated learning, with strong forgetting in the first store, and with slow or no forgetting in the second store. Here, on the one hand, the memory consolidation process of the second store favors a single burst at the start of the campaign, but the strong forgetting in the first store requires a uniform distribution of GRPs over the campaign weeks. A mixture of bursting and dripping also applies to one-store models with moderate forgetting. Examples of the above-mentioned brands and optimal GRP distributions are shown in Figure 5. Table 3 summarizes the above implications of the model for optimal advertisement scheduling.

Discussion

In this paper, we presented and applied a neurobiologically informed model to the recall of advertising material as a function of multiple exposures to the same advertisement (print or TV). To our knowledge, it is the first application of neurobiological process models to advertising. The same processes are involved in a wide range of learning and forgetting situations to which our model has been fitted successfully. This study shows that the model also captures the learning and forgetting of advertising material.

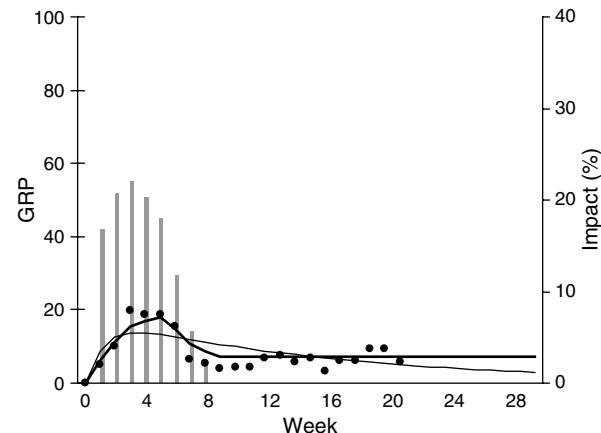
An important and essential feature of the model is its ability to describe situations with multiple learning trials, which was demonstrated in previous studies with model fits to psychological laboratory data, and in the present setting to recall data for multiple advertisement exposures. About 93% of the fitted recall-probability functions were not rejected. This shows that the recall functions, which are derived at neural and individual levels, give good approximations of the impact and brand name recall data at the aggregate level. Two reasons could explain this: (1) the variation in the number of advertisement exposures among respondents in an advertising week is rather small; (2) the memory processes

may be quite homogeneous over respondents. Departures from these properties will imply an overestimation of the recall functions used here at aggregate level, because recall probability is a concave function of memory intensity. Explicit modeling of variations over respondents should reveal the precise amount of overestimation. Because the overestimation applies to the entire course of the recall-probability function, we expect that the effects on the optimal schedule types shown in Table 3 will be limited.

The fits of our model to the SPOT data have shown that the feed-forward structure of the notion of memory stores, by which we allow memories to consolidate to one or more stores, is an essential element in the applicability of our model. There are numerous instances in the SPOT data that show a decline of impact and brand familiarity during advertising, which then stabilize at constant values after a campaign. From our model perspective, these cases all result from a rapid decline in the first memory store, from which a part of information is consolidated to a second store.

As a comparison, we implemented and applied the model proposed by Naik et al. (1998) to the SPOT data. Their model gave considerably poorer fits with as many free parameters as in the memory chain model, with about 22% of the fits rejected (Table 2). The poorer fits were caused to a large extent by the model rejections for situations described in the previous paragraph. The use of one forgetting parameter (δ in Naik et al. 1998) is not sufficient to capture complex behaviors of advertising effectiveness during and after campaigns. Naik et al.'s model must fit variations in memory decline during and after a campaign with the same value for the forgetting parameter δ . This leads to overly smooth fits, and an inability to

Figure 6 Fits of a Two-Store Memory Chain Model (Thick Line) and of the Model of Naik et al. (Thin Line), with the Same Number of Parameters, to Impact Data (Dots) of a Meal Sauce Brand



Notes. The model of Naik et al. is rejected at $\alpha = 0.05$, while the 2-store model is not rejected at very large test sizes ($\alpha > 0.85$).

handle changes in memory decline over time due to, for instance, long-term memory consolidation (see Figure 6). This is also true for other models, such as the Vidale-Wolfe and Nerlove-Arrow models and Brandaia (see the paper by Naik et al., where expressions for different models are listed in Table 2).

Our cross-validation study showed that the parameter estimates are stable over time. The models thus used to optimize advertisement scheduling showed that optimal schedules fall in one of three types: (1) a GRP burst at the start of the campaign, (2) a uniform, dripping GRP distribution, (3) a combination of bursting and dripping, with the greatest GRP expenditure at the start of the campaign and a more even distribution of GRP in later weeks. The implications of the memory model parameters for optimal scheduling, as shown in Table 3, could thus be used to target advertising to different groups of products or brands (see also Iyer et al. 2005).

The optimal schedules that we obtained are all decreasing or constant in GRP as a function of time. This does not mean, however, that the schedules implied by our model are all monotonic (e.g., decreasing). Further model analyses have shown that optimal GRP schedules may also be nonmonotonic. The critical parameters in this situation are the decline parameters of the memory stores and the rate of induction between the first and second stores. Nonmonotonic optimal schedules arise when the induction rate is larger than the memory decline rate of the first store. This situation never arose in our fits.

The contribution of this study from a scheduler's or manager's point of view is the effect of advertisement memory parameters on optimal scheduling. Decisions concerning the planning of advertisement campaigns could be supported better when knowledge is available about relations between memory parameters and advertisement characteristics. Although finding these relations is left as a challenge for future research, we could make several hypotheses here. For instance, we believe that a small saturation of learning parameter r_{\max} is related to a more or less worn-out advertising concept, that encoding μ_1 and learning rate v are highly influenced by personal involvement and attention, and that the recall threshold b is influenced by novelty and the coherence of an advertising message (i.e., small for well-known or coherent messages). Answers to these open questions (e.g., through psychological experiments) could lead to a better and more complete understanding of the relation

ad properties → memory parameters → scheduling.

In future research, we will extend our model with a decision-analytic component for consumer brand choice and buying behavior (e.g., see Akçura et al.

2004). We will also consider explicit modeling of competition effects. In this respect, it could be interesting to integrate the model part on copy wearout in Naik et al. (1998) with our model.

Acknowledgments

This research was supported by NWO, the Netherlands Society for Scientific Research. The authors are very grateful to Jan Lighthart for providing the SPOT data and to Giep Franzen for many helpful discussions. They also thank the reviewers for their many useful and detailed comments.

References

- Abeles, M. 1991. *Corticonics*. Cambridge University Press, Cambridge, UK.
- Akçura, M. T., F. F. Gönül, E. Petrova. 2004. Consumer learning and brand valuation: An application on over-the-counter drugs. *Marketing Sci.* **23** 156–169.
- Atkinson, R. C., R. M. Shiffrin. 1968. Human memory: A proposed system and its control processes. K. W. Spence, J. T. Spence, eds. *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 2. Academic Press, New York.
- Baddeley, A. D., D. J. A. Longman. 1978. The influence of length and frequency of training session on the rate of learning to type. *Ergonomics* **21** 627–635.
- Bahrück, H. P., L. E. Bahrück, A. S. Bahrück, P. E. Bahrück. 1993. Maintenance of foreign language vocabulary and the spacing effect. *Psych. Sci.* **4** 316–321.
- Barnes, C. A., B. L. McNaughton. 1980. Spatial memory and hippocampal synaptic plasticity. D. Stein, ed. *The Psychobiology of Aging: Problems and Perspectives*. Elsevier/North Holland, New York, 253–272.
- Burtt, H. E., E. M. Dobell. 1925. The curve of forgetting for advertising material. *J. Appl. Psych.* **9** 5–21.
- Chessa, A. G., J. M. J. Murre. 2004. A memory model for Internet hits after media exposure. *Physica A* **333** 541–552.
- Daley, D. J., D. Vere-Jones. 1988. *An Introduction to the Theory of Point Processes*. Springer-Verlag, Inc., New York.
- Diggle, P. J. 1983. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London, UK.
- Glenberg, A. M. 1976. Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *J. Verbal Learn. Verbal Behav.* **15** 1–16.
- Glenberg, A. M., T. S. Lehmann. 1980. Spacing repetitions over 1 week. *Memory Cognition* **8** 528–538.
- Goldman-Rakic, P. S. 1992. Working memory and the mind. *Sci. Amer.* **267** 111–117.
- Goldman-Rakic, P. S. 1995. Cellular basis of working memory. *Neuron* **14** 477–485.
- Iyer, G., D. Soberman, J. M. Villas-Boas. 2005. The targeting of advertising. *Marketing Sci.* **24** 461–476.
- Janssen, S. M. J., A. G. Chessa, J. M. J. Murre. 2005. The reminiscence bump in autobiographical memory: Effects of age, gender, education, and culture. *Memory* **13** 658–668.
- McGaugh, J. L. 2000. Memory—A century of consolidation. *Science* **287** 248–251.
- Meeter, M., J. M. J. Murre. 2004. Consolidation of long-term memory: Evidence and alternatives. *Psych. Bull.* **130** 843–857.
- Meeter, M., J. M. J. Murre. 2005. TraceLink: A model of consolidation and amnesia. *Cognitive Neuropsych.* **22** 559–587.
- Merzereau, P., L. Battaïs. 2000. Continuity or burst? How do TV strategies compare? *Admap* (October) 30–35.
- Murdock, B. B., Jr. 1961. The retention of individual items. *J. Experiment. Psych.* **62** 618–625.

- Murdock, B. B., Jr. 1974. *Human Memory: Theory and Data*. Lawrence Erlbaum Associates, Inc., Potomac, MD.
- Murre, J. M. J. 1996. TraceLink: A model of amnesia and consolidation of memory. *Hippocampus* **6** 675–684.
- Murre, J. M. J., K. S. Graham, J. R. Hodges. 2001. Semantic dementia: Relevance to connectionist models of long-term memory. *Brain* **124** 647–675.
- Murre, J. M. J., M. Meeter, A. G. Chessà. Modeling amnesia: Connectionist and mathematical approaches. M. J. Wenger, C. Schuster, eds. *Statistical and Process Models for Neuroscience and Aging*. Erlbaum, Mahwah, NJ. Forthcoming.
- Naik, P. A., M. K. Mantrala, A. G. Sawyer. 1998. Planning media schedules in the presence of dynamic advertising quality. *Marketing Sci.* **17** 214–235.
- Peterson, L. R., M. J. Peterson. 1959. Short-term retention of individual verbal items. *J. Experiment. Psych.* **58** 193–198.
- Rubin, D. C., A. E. Wenzel. 1996. One hundred years of forgetting: A quantitative description of retention. *Psych. Rev.* **103** 734–760.
- Rumelhart, D. E. 1967. The effects of interpresentation intervals on performance in a continuous paired-associate task. Technical Report 116, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, CA.
- Simon, J. L. 1979. What do Zielske's real data really show about pulsing? *J. Marketing Res.* **16** 415–420.
- SPOT. 1998. The effectiveness of TV-advertising. Internal report. (Available at www.spot.nl/onderzoek/onderzoek/downloads/Effectiviteit.pdf, in Dutch: *De effectiviteit van tv-reclame*), The Foundation for Promotion and Optimization of Television Advertising (SPOT), Amsterdam, The Netherlands.
- Stoyan, D., W. S. Kendall, J. Mecke. 1987. *Stochastic Geometry and Its Applications*. John Wiley Sons/Akademie Verlag, Berlin, Germany.
- Talamini, L. M., M. Meeter, J. M. J. Murre, B. Elvevåg, T. E. Goldberg. 2005. Integration of parallel input streams in parahippocampal model circuits. *Arch. General Psychiatry* **62** 485–493.
- Wickens, T. D. 1982. *Models for Behavior: Stochastic Processes in Psychology*. Freeman, San Francisco, CA.
- Zielske, H. A. 1959. The remembering and forgetting of advertising. *J. Marketing* **23** 239–243.

The Neural Mechanisms Underlying the Influence of Pavlovian Cues on Human Decision Making

Signe Bray,¹ Antonio Rangel,^{1,3} Shinsuke Shimojo,^{1,2} Bernard Balleine,⁴ and John P O'Doherty^{1,3}

¹Computation and Neural Systems, ²Division of Biology, and ³Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125, and ⁴Department of Psychology and the Brain Science Institute, University of California Los Angeles, Los Angeles, California 90024

In outcome-specific transfer, pavlovian cues that are predictive of specific outcomes bias action choice toward actions associated with those outcomes. This transfer occurs despite no explicit training of the instrumental actions in the presence of pavlovian cues. The neural substrates of this effect in humans are unknown. To address this, we scanned 23 human subjects with functional magnetic resonance imaging while they made choices between different liquid food rewards in the presence of pavlovian cues previously associated with one of these outcomes. We found behavioral evidence of outcome-specific transfer effects in our subjects, as well as differential blood oxygenation level-dependent activity in a region of ventrolateral putamen when subjects chose, respectively, actions consistent and inconsistent with the pavlovian-predicted outcome. Our results suggest that choosing an action incompatible with a pavlovian-predicted outcome might require the inhibition of feasible but nonselected action–outcome associations. The results of this study are relevant for understanding how marketing actions can affect consumer choice behavior as well as for how environmental cues can influence drug-seeking behavior in addiction.

Key words: instrumental conditioning; decision; fMRI; human; learning; pavlovian conditioning; reward

Introduction

It is well known that pavlovian cues associated with rewarding outcomes can exert biasing effects on action selection (Colwill and Rescorla, 1988; Balleine, 1994). A form of this effect relevant for decision making is outcome-specific transfer (Rescorla, 1994; Corbit et al., 2001; Holland, 2004; Corbit and Balleine, 2005; Corbit and Janak, 2007). In outcome-specific transfer, an animal's choice between multiple simultaneously available instrumental responses leading to different outcomes can be biased by the presentation of a pavlovian cue previously associated with one of those outcomes, such that the animal will tend to favor the instrumental action corresponding to the particular outcome with which that cue has been associated. Outcome-specific transfer effects are evident, for example, in the impact that in-store advertisements and other marketing strategies have on consumer behavior (Smeets and Barnes-Holmes, 2003), as well as in addictive behavior (Hogarth et al., 2007).

Lesion studies in rodents indicate that the following structures are necessary for outcome-specific transfer to occur: the striatum, including the nucleus accumbens shell (Corbit et al., 2001) and the dorsolateral striatum (Corbit and Janak, 2007), and structures afferent to these regions, including the mediolateral orbito-

frontal cortex (Ostlund and Balleine, 2007) and basolateral amygdala (Corbit and Balleine, 2005).

Outcome-specific transfer can be differentiated from another form of pavlovian-instrumental interaction called general transfer, in which a pavlovian cue exerts a nonspecific energizing effect on instrumental behavior by increasing the vigor of instrumental responses (Holland, 2004; Corbit and Balleine, 2005). General transfer seems to depend on circuitry involving the ventral striatum and amygdala that is clearly dissociable from that involved in the outcome-specific transfer effect: lesions of the nucleus accumbens core and amygdala central nucleus affect general transfer but leave specific transfer intact, whereas lesions in the nucleus accumbens shell and basolateral amygdala have the converse effect (Corbit et al., 2001; Corbit and Balleine, 2005). In humans, a recent functional magnetic resonance imaging (fMRI) study has implicated human nucleus accumbens and amygdala in general transfer (Talmi et al., 2008), but the brain systems underlying outcome-specific transfer in the human or primate brain more generally have yet to be identified. Furthermore, whereas rodent lesion studies have identified regions that appear to be necessary for specific transfer, the precise functional contribution of each of these regions to this process has yet to be characterized.

The aim of the present study was twofold: first, to determine the neural substrates of the outcome-specific transfer effect in the human brain, and, second, to gain insight into the neural computations within these regions that might underlie this function. To address these aims, we used event-related fMRI to measure blood oxygenation level-dependent (BOLD) responses in human subjects while they made instrumental choices in the presence of pavlovian cues that were either associated with the liquid food reward outcomes generated by some of the actions, or associated

Received Feb. 28, 2008; revised April 17, 2008; accepted May 5, 2008.

This work was supported by a Searle scholarship and grants from the National Science Foundation (NSF061714) (J.O.D.), the Gordon and Betty Moore Foundation (J.O.D., A.R.), and the Caltech Brain Imaging Center and the Japan Science and Technology Agency (ERATO) (S.S.). We are grateful to Vivian Valentin for assistance and to Hackjin Kim, Sam Huang, and Shawn Wagner for developing the coil we used to detect swallowing movement.

Correspondence should be addressed to John P O'Doherty, Division of the Humanities and Social Sciences, California Institute of Technology, MC 228-77, Pasadena, CA 91125. E-mail: jdoherty@hss.caltech.edu.

DOI:10.1523/JNEUROSCI.0897-08.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/285861-06\$15.00/0

with an affectively neutral (control) outcome. On the basis of the animal studies, we focused our analysis on the striatum, particularly its ventral aspect, including the nucleus accumbens and adjacent ventral putamen. We also tested for outcome-specific transfer effects in the amygdala.

Materials and Methods

Subjects

Twenty-three healthy, right-handed subjects participated in this study (six females), ranging in age from 18 to 40 (mean 24 ± 5.3 SD). One additional subject did not complete the study and was not included in the analysis. All subjects gave informed consent and the study was approved by the Caltech Institutional Review Board.

Stimuli

Visual stimuli were presented via a projector positioned at the back of the room. Subjects viewed a reflection of the projected image (800×600 pixels) in a mirror attached to the scanner head coil. The food rewards were delivered by means of four separate electronic syringe pumps (one for each liquid) positioned in the scanner control room. For each rewarded trial, these pumps pushed 0.6 ml of liquid to the subject's mouth via ~ 10 m long polyethylene plastic tubes, the other end of which were held between the subject's lips like a straw while they lay supine in the scanner. Stimulus presentation and response recording were controlled with the Cogent 2000 Matlab (Mathworks) toolbox.

Behavioral procedures

During both the pavlovian and instrumental training subjects were explicitly asked to learn the cue–outcome and action–outcome relationships. All four training and test sessions described below were performed in the scanner.

Pavlovian training. Pavlovian training consisted of the presentation of associations between simple geometrical visual stimuli (Fig. 1*a*) and one of four liquid outcomes, three of which were rewarding [chocolate milk (Hershey's, distributed by Dean National Brand Group), cola (Coca-Cola), and orange juice (Trader Joe's)] and an affectively neutral tasteless control solution, which consisted of the main ionic components of human saliva (25 mM KCl and 2.5 mM NaHCO₃) (Fig. 1*a*). Cues were presented at the center of the screen for 1.75 s, and then 3 s after cue offset rewards were delivered with a probability of 50%. The intertrial interval varied uniformly between 1 and 5 s.

Instrumental training. During instrumental training trials, subjects were asked to choose between two button-press actions. Four gray squares at the bottom of the screen corresponded to the four buttons on a response box (Current Designs) that the subjects held in their right hand. Specific actions were made available for selection when the corresponding squares changed color from gray to brown, two at a time. As in the pavlovian trials, the response cues appeared for 1.75 s. Subjects were asked to make a choice during this time. The choice was followed by a 3 s delay before the outcome associated with the chosen action was delivered on 50% of trials (Fig. 1*b*). The intertrial interval varied uniformly between 1 and 5 s. Responses on each button earned distinct outcomes: two of the buttons led to rewarding outcomes (for example, orange juice and chocolate milk), and two led to the neutral outcome. Therefore, during pavlovian training, subjects experienced four different outcomes, whereas in the instrumental trials they experienced only three.

Training schedule. The first training session consisted entirely of pavlovian trials, 10 of each type for a total of 40 trials and a duration of ~ 6 min (Table 1). The second training session consisted entirely of instrumental trials, six of each type for a total of 36 trials and a duration of ~ 5 min. In the first two sessions, pavlovian and instrumental trials were presented separately to enhance learning of the respective associations. In the third session, pavlovian and instrumental trials were randomly intermixed, 60 (15 \times 4) pavlovian trials and 60 (10 \times 6) instrumental trials, for a duration of ~ 18 min. Before training and after each session, subjects rated the pleasantness of the stimuli as described below.

Outcome-specific transfer. After the three training sessions subjects performed a series of transfer trials. During transfer trials one of the pavlovian cues was presented simultaneously with the instrumental cues (Fig.

Table 1. Trial composition for training and transfer sessions

Phase	Number of presentations	Cues			Outcome
		Pavlovian	Instrumental		
1	10	S1			01
		S2			02
		S3			03
		S4			ON
2	6		R1	R2	01
			R1	R3	01
			R1	R4	01
			R2	R3	02
			R2	R4	02
			R3	R4	ON
3	10	S1			01
		S2			02
		S3			03
		S4			ON
			R1	R2	01
			R1	R3	01
			R1	R4	01
			R2	R3	02
			R2	R4	02
			R3	R4	ON
4	25	S1	R1	R2	02
		S2	R1	R2	
		S3	R1	R2	
		S4	R1	R2	
		S4	R3	R4	

S1–S4, Visual cues; R1–R4, four button response actions; 01–03, liquid rewards; ON, affectively neutral, tasteless control solution.

1*c*), and as in instrumental training, subjects were asked to choose between two available options. This phase was conducted in extinction, meaning that no outcomes were delivered. The reason for performing this phase in extinction was to allow assessment of the influence of the pavlovian cues on instrumental responding without the confounding effects of the outcomes themselves. Testing for outcome-specific effects in extinction is standard in animal learning studies of this phenomenon (Rescorla, 1994; Blundell et al., 2001; Corbit et al., 2001).

There were five different types of trials. Two of the trial types were designed to test for outcome-specific transfer effects. On these trials, subjects chose between actions associated with two particular reward outcomes, O1 and O2 (for example, orange juice and chocolate milk), whereas the concurrently presented pavlovian cue was associated with one of these specific outcomes. One of the specific trial types involved the pavlovian cue paired with outcome O1, and the other specific trial type involved the pavlovian cue paired with O2. Evidence for an outcome-specific transfer effect would be seen if the presence of the pavlovian cue biased choice toward the action associated with the same outcome as the pavlovian cue. In the subsequent analysis, we pooled over both of the specific trial types, but differentiated between trials in which subjects made choices compatible with the pavlovian outcome from those trials in which subjects made choices that were not compatible.

In a third “pavlovian reward control” trial type, subjects were again presented with the choice between two reward outcomes (O1 and O2), but instead the pavlovian cue was previously associated with a different outcome (for example, cola), that was not compatible with either response option.

In the fourth “pavlovian neutral control” trial type, subjects were again presented with the choice between two reward outcomes (O1 and O2), but the pavlovian cue presented this time was that associated with the affectively neutral outcome.

In the final “neutral choice control” trial type, subjects made choices between actions associated with the affectively neutral outcome, in the presence of a pavlovian cue also associated with a neutral outcome. This last trial type was intended to be a baseline condition for choosing between two options in the presence of a visual cue but in the absence of predicted rewards. Each type of trial was presented 25 times, for a total of 125 trials and a duration of ~ 20 min.

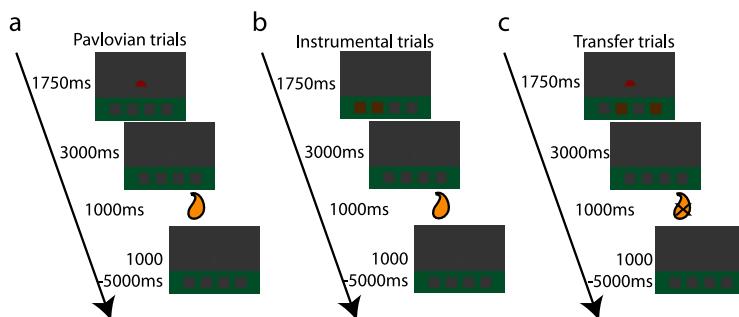


Figure 1. Illustration of the three different trial types in this study. **a**, Pavlovian trial. A visual shape stimulus was presented at the center of the screen for 1.75 s followed by a fixation cross for 3 s. The liquid outcome corresponding to the stimulus was then delivered with a probability of 50%. One second was allotted for consumption, and the interval between trials varied uniformly between 1 and 5 s. **b**, Instrumental trial. Two of the four squares at the bottom of the screen changed color from gray to brown for 1.75 s during which time subjects were instructed to push one of the buttons. The liquid outcome corresponding to their response was delivered after 3 s, with a probability of 50%. One second was allotted for consumption, and the interval between trials varied uniformly between 1 and 5 s. **c**, Transfer trials. A visual shape stimulus was presented simultaneously with two squares changing color. Subjects were instructed to press one of the corresponding buttons. Timing was similar to the pavlovian and instrumental trials; however, no outcomes were delivered during these trials.

Behavioral measures

Reaction times. Reaction times to choices were recorded both during the learning trials and the transfer test trials; these can be used as an online measure of learning (O'Doherty et al., 2006).

Pupillary dilation. Pupil diameter was continuously measured during scanning using an Applied Science Laboratories MRI-compatible eye-tracking system. Pupil reflex amplitude has been shown to be modulated by arousal level and can thus be used as an index of conditioning (Bitsios et al., 2004).

Affective evaluations of stimuli. Before the start of the training procedure, and after each scanning session, we asked subjects to rate the pleasantness of the shape images and the liquid outcomes. Within each category, stimuli were presented in random order and subjects reported their evaluation by moving a cursor along a scale from -5 to $+5$.

Swallowing motion. A motion-sensitive inductive coil was positioned on top of the subjects' throat using a Velcro strap around the neck. This measured the motion of the subjects' throat during swallowing. The time course derived from this measure was used as a regressor of no interest in the fMRI data analysis. We do not have recordings for one subject who found the coil uncomfortable.

fMRI scanning procedure

fMRI data were acquired on a Siemens 3T TRIO MRI scanner; BOLD contrast was measured with gradient echo T2*-weighted echo-planar images (EPIs). Imaging parameters were optimized to minimize signal dropout in medial ventral prefrontal and anterior ventral striatum: we used a tilted acquisition sequence at 30° to the anterior cingulate-posterior cingulate line (Deichmann et al., 2003), and an eight-channel phased array coil, which yields a $\sim 40\%$ signal increase in this area over a standard coil. The first five volumes of each session were discarded to permit T1 equilibration. Other parameters were as follows: 36 slices, in-plane resolution, 3×3 mm; slice thickness, 3 mm; repetition time, 2.25 s; echo time, 30 ms; field of view, 192×192 mm. A T1-weighted structural image was also acquired for each subject, as well as a 49-slice EPI to improve coregistration.

Imaging data processing and analysis

Data were preprocessed using the SPM5 software package (<http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>). Images were corrected for slice timing and spatially realigned to the first image from the first session. One of the 49-slice EPIs collected at the end of the experiment was used to improve coregistration and spatial normalization. The 36-slice EPIs were coregistered to a 49-slice EPI, which was in turn coregistered to the T1-weighted anatomical scan. The T1 image was segmented into white and gray matter, and the gray matter was coregistered and normalized to the template gray matter image distributed with SPM5 (in Montreal

Neurological Institute space). These parameters were subsequently applied to the T1 image itself as well as the set of 36-slice EPIs. Spatial smoothing was then applied to the 36-slice EPIs using a Gaussian kernel with a full width at half maximum of 8 mm.

Statistical analysis was performed using a general linear model (GLM). The transfer session was modeled separately from the three training sessions, and here we report results only from the transfer phase of the experiment. The GLM included regressors at the time of cue onset for five conditions: specific transfer when the option compatible with the pavlovian cue was chosen, specific transfer when the incompatible option was chosen, pavlovian reward control, pavlovian neutral control, and neutral choice control. We also included regressors at the time of expected outcome. Each regressor was modeled as an impulse function (0 s), and convolved with the canonical hemodynamic response function. Regressors of no interest included missed trials when no option was chosen, the six ongoing motion parameters estimated during realignment, and motion caused by swallowing. The results from each subject were taken to the random effects level by applying *t* tests between contrast images to produce group statistical parametric maps.

Results

Behavioral results

Results of pavlovian training

Behavioral results indicate that the pavlovian stimulus–outcome associations were successfully learned. After each training session, subjects were asked to rate on a scale from -5 to $+5$ how pleasant they found each shape stimulus and each liquid. After training, subjects rated the stimuli associated with rewarding outcomes as significantly more pleasant than the stimulus associated with the neutral outcome (paired *t* test, $t_{(22)} = -3.0840$; $p < 0.01$) (Fig. 2a). Pupil reflex amplitude also discriminated between reward and neutral conditions (Fig. 2b). In the 16 subjects who showed reliable amplitude changes in pupil diameter after cue presentation, the peak amplitude is significantly smaller for rewarded outcome trials, which indicates a higher degree of arousal when subjects saw reward predictive cues (paired *t* test, $t_{(15)} = 2.4173$; $p < 0.05$) (Bitsios et al., 2004; Seymour et al., 2007).

Initial learning of instrumental associations

Subjects' choice behavior in the instrumental trials indicated that the instrumental associations were acquired. During the final training session subjects were significantly more likely to choose the action delivering a reward outcome when the alternative action delivered the neutral solution (Fig. 2c) (one-sided paired *t* test, $t_{(22)} = 1.8399$; $p < 0.05$).

Outcome-selective transfer effects during test phase

We found evidence for an outcome-specific transfer effect in subjects' choice behavior during the transfer test phase. During the transfer phase, subjects choose the compatible option on average 66% of the time; this is significantly higher than cue invariant responding which averages to 50% over the two outcome-specific conditions (paired *t* test, $t_{(22)} = 3.6348$; $p < 0.005$). There were a total of 50 specific transfer trials for each subject and, separating these into five 10-trial bins, we found that there was neither a significant increase or decrease in choice allocation across time

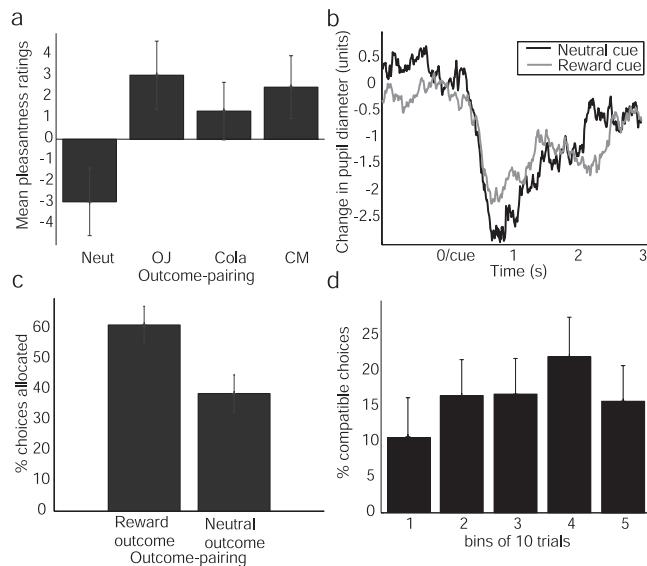


Figure 2. Behavior during training and test sessions. **a**, Mean pleasantness ratings for visual cue stimuli after the training sessions, plotted by outcome pairing. The cues paired with the neutral outcome were rated as significantly less pleasant than the cues paired with reward outcomes (paired t test, $t_{(22)} = -3.0840$; $p < 0.01$). **b**, Pupil diameter in response to visual cues. The peak amplitude is significantly smaller for the cues paired with reward outcomes for the 16 subjects who showed reliable amplitude changes after cue presentation (paired t test, $t_{(15)} = 2.4173$; $p < 0.05$). **c**, Choice behavior during the second session of instrumental trials, above cue invariant responding (50%). Plotted are responses during trials in which subjects chose between a reward outcome and the neutral outcome. Subjects were significantly more likely to choose the action leading to the reward outcome (one-sided paired t test, $t_{(22)} = 1.8399$; $p < 0.05$). **d**, Choice data binned into five 10-trial bins. There is no significant linear trend across the session (linear regression of the percentage of compatible choice allocation onto bin number, $p = 0.239$). Error bars indicate SEM.

(Fig. 2d), indicating that the biasing effect of the pavlovian cues on choice persisted for the duration of the extinction test and did not attenuate.

fMRI results

To gain insight into the mechanisms underlying outcome-specific transfer in humans, we performed two analyses. First, we compared brain activity during trials assessing outcome-specific transfer when subjects chose the option compatible with the pavlovian cue to trials when they chose the incompatible option (one subject who never chose the incompatible cue was excluded from this analysis) (Fig. 3a). We found significant activation in the right ventrolateral putamen ($t_{(21)} = 3.79$, $p < 0.001$ uncorrected; $x = 27$, $y = -3$, $z = -3$) extending posteriorly toward the pallidum ($t_{(21)} = 3.81$, $p < 0.001$ uncorrected; $x = 24$, $y = -18$, $z = 0$). The left pallidum also showed a peak at this threshold ($t_{(21)} = 3.82$, $p < 0.001$ uncorrected; $x = -27$, $y = -15$, $z = -3$). These were the only regions to meet our significance criterion in this contrast.

Second, we plotted the average parameter estimates taken from the general linear model estimates at the peak putamen voxel for each subject (Fig. 3b). We found that the difference between conditions was caused by a significant decrease in signal during the outcome-specific trials where the incompatible response was chosen, relative to the outcome-specific trials when the compatible response was chosen and to the other control conditions. In fact, activity in the compatible condition did not differ significantly from activity during any of the other control conditions (paired t tests, $p > 0.05$) and, more generally, activity

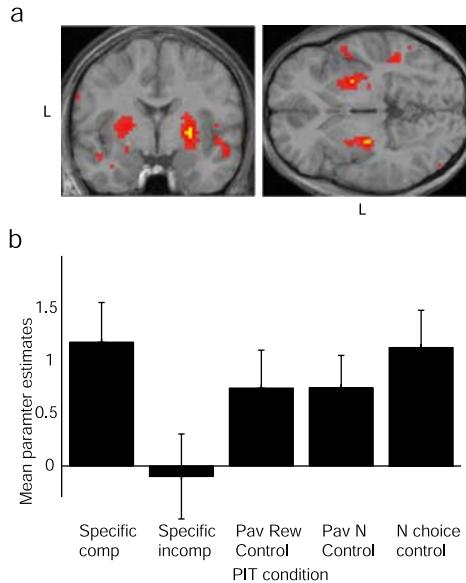


Figure 3. Imaging results from the pavlovian-instrumental transfer (PIT) phase. **a**, fMRI results from the contrast comparing the outcome-specific transfer trials in which the action compatible with the pavlovian cue is selected to those in which the incompatible action is selected (red, $p < 0.01$; yellow, $p < 0.001$). At a threshold of $p < 0.001$, uncorrected, we find significant activation in the ventrolateral putamen ($t_{(21)} = 3.79$, $p < 0.001$, uncorrected; $x = 27$, $y = -3$, $z = -3$) and bilateral pallidum ($t_{(21)} = 3.81$, $p < 0.001$, uncorrected, $x = 24$, $y = -18$, $z = 0$; and $t_{(21)} = 3.82$, $p < 0.001$, uncorrected, $x = -27$, $y = -15$, $z = -3$). **b**, Parameter estimates from the peak putamen voxel for each subject for each of the five experimental conditions during the transfer phase [specific compatible (comp), specific incompatible (incomp), pavlovian reward (pav rew) control, pavlovian neutral (pav N) control, and neutral (N) choice control]. Parameter estimates in the specific compatible condition do not differ significantly from any condition other than specific incompatible (paired t tests, $p > 0.05$). Error bars indicate SEM.

in the outcome-specific trials did not differ from the control conditions (paired t tests, $p > 0.05$).

Discussion

Our results provide insights into the neural mechanisms by which pavlovian cues can modulate choice between different instrumental courses of action, in humans. In outcome-specific transfer, subjects are more likely to choose an action that is associated with a particular outcome in the presence of a pavlovian cue that was previously associated with the presence of that outcome. We found neural correlates of outcome-specific transfer in a very circumscribed region of extended ventral striatum in the ventral caudolateral putamen. This region and an adjacent region of ventral pallidum were the only areas to meet our statistical criterion for significance.

These findings add to an accumulating body of evidence from human fMRI studies of a role for an extended region of ventral parts of putamen alongside the nucleus accumbens in functions related to reward-learning and prediction errors (O'Doherty et al., 2002, 2003; McClure et al., 2003) and now in interactions between pavlovian and instrumental conditioning. Such findings resonate with anatomical and histochemical studies in primates that indicate that ventral parts of putamen share many of the cytoarchitectonic characteristics of nucleus accumbens, as well as sharing similar inputs (Russchen et al., 1985; Selemon and Goldman-Rakic, 1985; Fudge and Haber, 2002; Fudge et al., 2002).

The present findings do suggest, however, that different parts of the ventral striatum may contribute differentially to distinct forms of pavlovian-instrumental transfer in humans. This sug-

gestion is based on a comparison of our finding that the ventrolateral putamen is involved in outcome-specific transfer in humans with the results of a previous study implicating nucleus accumbens in the general excitatory effects of pavlovian cues on instrumental performance (Talmi et al., 2008). It is well established that pavlovian cues can exert a general, nonspecific excitatory effect on the performance of instrumental actions (Estes, 1943, 1948; Rescorla and Solomon, 1967; Colwill and Rescorla, 1988; Rescorla, 1994; Holland, 2004), an effect that Talmi et al. (2008) demonstrated is mediated by activation of nucleus accumbens and of amygdala. In the context of the present study, these findings suggest that outcome-specific and general transfer may depend on quite distinct neural substrates in humans, mirroring clear double dissociations between the neural circuits known to be involved in implementing these effects in rodents (Corbit et al., 2001; Corbit and Balleine, 2005). Although the present study was not designed to assess the effects of general transfer, in future it will be important to compare and contrast outcome-specific and general transfer effects within the same fMRI study to provide a more direct test of the hypothesis that, as in rodents, outcome-specific and general transfer in humans depend on distinct components of the ventral striatum.

Note that although we found a remarkably good correspondence between our findings and those from the rodent lesion studies at the level of the ventral striatum, other regions in addition to the ventral striatum have been implicated in specific pavlovian-instrumental transfer in rodents, including the basolateral amygdala (Corbit and Balleine, 2005) and dorsolateral striatum (Corbit and Janak, 2007). We did not find any evidence for a differential contribution of these regions in the present study. One possibility is that these areas do play a role in specific transfer effects in humans, but this does not result in a global increase in activity between conditions and, thus, does not become manifest with BOLD fMRI.

The present results go beyond merely pointing to homologies between outcome-specific transfer effects in rodents and humans. Previous animal studies on this topic have all involved lesion manipulations, which, although important for identifying whether a given region is necessary for implementing specific transfer effects, cannot provide insight into the neural computations underlying such an effect. Here, we measured dynamic changes in BOLD responses as subjects made choices that were either consistent or inconsistent with the specific transfer effect. Responses consistent with the specific transfer effect occurred when subjects chose the outcome compatible with the pavlovian cue, and inconsistent responses occurred when subjects chose the incompatible option. Although subjects showed a significant bias toward the compatible action overall, sometimes they chose the incompatible action; this allowed us to compare activity when transfer guided behavior with activity under identical stimulus conditions when subjects chose independently of the cue. Activity in ventrolateral putamen was not significantly elevated on trials when an outcome-specific cue was presented compared with control trials where cues for other, unavailable outcomes were presented, suggesting that outcome-specific transfer effects are not mediated by an overall increase in activity in this area. Furthermore, even on outcome-specific trials where subjects chose the action compatible with the pavlovian cue, there was no increase in activity compared with non-outcome-specific control trials. Instead, we found a significant decrease in signal on those outcome-specific trials where subjects chose the action incompatible with the outcome, compared with compatible choice outcome-specific trials.

This finding provides insight into the computations that might be taking place in the ventral striatum during outcome-specific transfer effects. Outcome-specific transfer effects are thought to be mediated by outcome–response (O–R) associations that are activated by the pavlovian cues (Rescorla, 1994; Balleine and Ostlund, 2007). A natural hypothesis is that when the action plan activated by the O–R association is feasible (because such an action is available), it must be inhibited before another action can be taken. Note that, under this hypothesis, the O–R association needs to be inhibited during the outcome-specific transfer trials when the incompatible response is chosen, but not when the compatible response is selected, or in any of the other control trials. This provides a computational explanation for why suppression of activity in the ventrolateral putamen is observed only in the incompatible outcome-specific transfer trials.

Specific transfer effects from pavlovian cues have been argued to play a role in addictive behaviors (Ludwig et al., 1974). For example, Hogarth et al. (2007) demonstrated specific transfer of a tobacco-seeking response in the presence of a tobacco predicting cue, relative to a money predicting cue. Here, we demonstrate similar behavioral results, using nonaddictive outcomes, indicating that the observed transfer effects reflect a general property of reward-associated cues that are not specifically related to addictive stimuli. Nonetheless, there are clear parallels between our experimental design and the potential influence of environmental cues on drug-seeking behavior. Our fMRI results suggest the hypothesis that suppression of an outcome–response association might contribute toward biasing behavior away from cue-compatible responding. This raises the possibility of a future therapeutic intervention in addiction, in which ventrolateral putamen circuitry could potentially be targeted (for instance via a neurofeedback procedure) (deCharms et al., 2005; Bray et al., 2007) to suppress effects of environmental drug cues on drug-seeking behavior.

In this study, we demonstrated an outcome-specific pavlovian-instrumental transfer effect in humans, which serves to bias action choice toward actions associated with an outcome consistent with a concurrently presented cue. BOLD fMRI measured while subjects performed this task demonstrated a signal decrease in ventrolateral putamen when subjects' chose the action incompatible with the cue. This finding points to a computational role for this region in suppressing outcome–response associations, necessary to perform an action incompatible with the pavlovian cue only when a compatible action is feasible. This work adds to our understanding of the neural mechanisms of stimulus–outcome guided decision making in both animals and humans, which is fundamental for understanding maladaptive choice behaviors such as addiction.

References

- Balleine B (1994) Asymmetrical interactions between thirst and hunger in pavlovian-instrumental transfer. *Quart J Exp Psychol B* 47:211–231.
- Balleine BW, Ostlund SB (2007) Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann NY Acad Sci* 1104:147–171.
- Bitsios P, Szabadi E, Bradshaw CM (2004) The fear-inhibited light reflex: importance of the anticipation of an aversive event. *Int J Psychophysiol* 52:87–95.
- Blundell P, Hall G, Killcross S (2001) Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *J Neurosci* 21:9018–9026.
- Bray S, Shimojo S, O'Doherty JP (2007) Direct instrumental conditioning of neural activity using functional magnetic resonance imaging-derived reward feedback. *J Neurosci* 27:7498–7507.

- Colwill RM, Rescorla RA (1988) Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J Exp Psychol Anim Behav Process* 14:155–164.
- Corbit LH, Balleine BW (2005) Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci* 25:962–970.
- Corbit LH, Janak PH (2007) Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of pavlovian stimuli on instrumental responding. *J Neurosci* 27:13977–13981.
- Corbit LH, Muir JL, Balleine BW (2001) The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J Neurosci* 21:3251–3260.
- deCharms RC, Maeda F, Glover GH, Ludlow D, Pauly JM, Soneji D, Gabrieli JD, Mackey SC (2005) Control over brain activation and pain learned by using real-time functional MRI. *Proc Natl Acad Sci USA* 102:18626–18631.
- Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage* 19:430–441.
- Estes WK (1943) Discriminative conditioning. I. A discriminative property of conditioned anticipation. *J Exp Psychol* 32:150–155.
- Estes WK (1948) Discriminative conditioning. 2. Effects of a pavlovian conditioned stimulus upon a subsequently established operant response. *J Exp Psychol* 38:173–177.
- Fudge JL, Haber SN (2002) Defining the caudal ventral striatum in primates: cellular and histochemical features. *J Neurosci* 22:10078–10082.
- Fudge JL, Kunishio K, Walsh P, Richard C, Haber SN (2002) Amygdaloid projections to ventromedial striatal subterritories in the primate. *Neuroscience* 110:257–275.
- Hogarth L, Dickinson A, Wright A, Kouvaraki M, Duka T (2007) The role of drug expectancy in the control of human drug seeking. *J Exp Psychol Anim Behav Process* 33:484–496.
- Holland PC (2004) Relations between pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process* 30:104–117.
- Ludwig AM, Wikler A, Stark LH (1974) First drink: psychobiological aspects of craving. *Arch Gen Psychiatry* 30:539–547.
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- O'Doherty JP, Deichmann R, Critchley HD, Dolan RJ (2002) Neural responses during anticipation of a primary taste reward. *Neuron* 33:815–826.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ (2006) Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron* 49:157–166.
- Ostlund SB, Balleine BW (2007) Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome. *Learn Behav* 35:43–52.
- Rescorla RA (1994) Transfer of instrumental control mediated by a devalued outcome. *Anim Learn Behav* 22:27–33.
- Rescorla RA, Solomon RL (1967) Two-process learning theory: relationships between pavlovian conditioning and instrumental learning. *Psychol Rev* 74:151–182.
- Russchen FT, Bakst I, Amaral DG, Price JL (1985) The amygdalostriatal projections in the monkey. An anterograde tracing study. *Brain Res* 329:241–257.
- Seelen LD, Goldmanrakic PS (1985) Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *J Neurosci* 5:776–794.
- Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007) Differential encoding of losses and gains in the human striatum. *J Neurosci* 27:4826–4831.
- Smeets PM, Barnes-Holmes D (2003) Children's emergent preferences for soft drinks: stimulus-equivalence and transfer. *J Econ Psychol* 24:603–618.
- Talmi D, Seymour B, Dayan P, Dolan RJ (2008) Human pavlovian-instrumental transfer. *J Neurosci* 28:360–368.

Mere Exposure and the Endowment Effect on Consumer Decision Making

GAIL TOM

CAROLYN NELSON

TAMARA SRZENTIC

RYAN KING

California State University, Sacramento

ABSTRACT. Previous researchers (e.g., J. A. Bargh, 1992, 2002) demonstrated the importance of nonconscious processes on consumer choice behavior. Using an advertisement, the authors determined the effect of two nonconscious processes—the *mere exposure effect*, which increases object preference by increasing consumer exposure to an object, and the *endowment effect*, which increases object valuation by providing consumer possession of an object—on consumer behavior. Although the mere exposure effect and endowment effect did not produce an interaction, they produced independent effects. The endowment effect increased object valuation but not object preference. The mere exposure effect increased object preference but not object valuation. Thus, at the unconscious level, an increase in object preference does not lead to an increase in object valuation, nor does an increase in object valuation lead to an increase in object preference. The authors discuss the importance of developing measures of unconscious process in advertising effectiveness.

Keywords: consumer decision making, endowment effect, mere exposure, prime

RESEARCHERS IN CONSUMER DECISION MAKING have traditionally adopted a cognitive approach, assuming that consumer choice is a conscious and deliberate process (Cohen & Chakravarti, 1990; Jacoby, Johar, & Morrin, 1998; Simonson, Carmon, Dhar, Drolet, & Nowlis, 2001). However, researchers in social behavior and goal pursuit have increasingly called attention to the importance of nonconscious processes on consumer choice (Bargh, 2002; Fitzsimons, Hutchinson, & Williams, 2002). Consumer information processing and decision making are affected by motives that can be activated by nonconscious primes delivered subliminally, without the recipient's awareness, or supraliminally, with the recipient's awareness of the prime but not its intent to influence (Bargh,

This project was funded by a grant from the College of Business Administration, California State University, Sacramento.

Address correspondence to Gail Tom, College of Business Administration, California State University, Sacramento, CA 95819; tomgk@csus.edu (e-mail).

1992). Consumers choose products with characteristics that match their primed-activated motives. In one study, thirst-induced consumers subliminally exposed (i.e., primed) to happy facial expressions subsequently consumed more and were willing to pay more for a drink as compared with thirst-induced consumers who were subliminally exposed to angry faces (Winkielman, Berridge, & Wilbarger, 2001). Strahan, Spencer, and Zanna (2001) found that only consumers who were made to be thirsty and subliminally primed to be thirsty, but not nonthirsty, subliminally primed consumers, opted to drink more of a purported thirst-quenching drink than of a purported energy drink. Participants were not aware of the primed-activated motive and its subsequent effect on choice behaviors in either study. Results of these and other studies (Bargh, Chen, & Burrows, 1996; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Troetschel, 2001; Chartrand & Bargh, 2002; Fitzsimons & Shiv, 2001) demonstrated that consumer downstream behaviors are affected by previous unconscious processes and underscored the importance of nonconscious processes in consumer evaluation and choice behavior.

In our study, we extended the findings of previous researchers and investigated the interactive nature of nonconscious processes. We investigated anecdotal and empirical evidence supporting the relationship between object valuation and object preference, not only because of its face validity, but also because of its frequent occurrence. Consumers willingly pay more for a brand that they prefer than for comparable alternatives. We explored whether nonconscious-caused increases in participants' object preferences cause nonconscious-caused increases or decreases in participants' object valuation.

The *mere exposure effect* occurs when repeated or single exposure to a stimulus, even in the absence of awareness, results in the formation of a positive affective reaction to the stimulus (Zajonc, 1968). It is a simple way to increase object preference. Results of decades of research have demonstrated the robustness of the mere exposure phenomenon with a variety of stimulus domains and levels of participant awareness (Bornstein & D'Agostino, 1992). The mere exposure phenomenon cannot be explained by an appeal to recognition memory, perceptual fluency (Zajonc, 2001), or subjective familiarity (Wilson, 1979). It suggests that affect and cognition, although they contribute jointly to behavior, are separate psychological processes that can be influenced independently of one another (i.e., preference can be formed without cognitive mediation; Zajonc, 1968). LeDoux (1996) and Zola-Morgan, Squire, Alvarez-Royo, and Clower (1991) found neurophysiological support for independent affect and cognition. On the basis of results suggesting that stimuli perceived without awareness produce the mere exposure effect (Bornstein, 1989; Bornstein & D'Agostino), we attempted to create a methodology to produce nonconscious preference manipulation.

The *endowment effect* on object valuation is similar to the mere exposure effect on object preference (Thaler, 1980). The endowment effect refers to the finding that consumers assign more value to objects simply because they own them. The endowment effect has been demonstrated repeatedly and robustly with

a variety of products, such as binoculars (Kahnemann, Knetsch, & Thaler, 1990), pens (Knetsch, 1990), mugs (Tom, 2004), and candy bars (Knetsch, 1989), and in varied settings, such as laboratories (Kahneman et al.), field studies (Johnson, Hershey, Meszaros, & Kunreuther, 1993), and economic market experiments (Franciosi & Kujal, 1996).

The most accepted theoretical explanation for the endowment effect is loss aversion (Kahneman & Tversky, 1984). Aversion to the loss of a possession is greater than the gain of the possession. The endowment effect occurs outside of conscious awareness; people are not aware that ownership results in their greater valuation of the object. The endowment effect thus provides a technique to manipulate nonconscious object valuation.

The impacts of the mere exposure effect and the endowment effect are on opposite sides of the affect–cognition relationship. Whereas the mere exposure effect increases object preference (affect) without cognitive mediation, the endowment effect increases object valuation (cognitive) without affect mediation. Researchers studying the endowment effect found that ownership adds value to an object without consumers' awareness of its effect (e.g., Franciosi & Kujal, 1996; Johnson et al., 1993; Tom, 2004), and researchers studying the mere exposure effect found that exposure without awareness produced greater preference (e.g., Bornstein, 1989; Bornstein & D'Agostino, 1992; Zajonc, 1968, 2001). The endowment effect increases object valuation as a result of possession and may or may not result in a concomitant increase in object preference. In a similar relationship, the mere exposure effect increases object preference with or without a concomitant increase in object valuation.

On the basis of this research, we formed four hypotheses (H). If the endowment effect and the mere exposure effect are separate, independent processes, then the mere exposure effect will lead to greater object preference, but not to greater object valuation (H_1), and the endowment effect will lead to greater object valuation, but not to greater object preference (H_2). If the endowment and mere exposure effects interact, then the mere exposure effect and the endowment effect together will produce a greater object preference (H_3) and object valuation (H_4) than either effects produce alone.

Method

Materials

We investigated the effect of the endowment effect and mere exposure effect on object valuation and object preference with a 2 (Ownership: endowment vs. nonendowment) \times 3 (Exposure: no exposure, normal exposure, subliminal exposure) factorial design. In the endowed condition, we gave participants an object to examine and keep. In the unendowed condition, we gave participants an object to examine and then return. We created three versions of the same video for the

exposure condition. The subliminal condition delivered a nonconscious exposure. The normal, or supraliminal, condition provided an exposure representative of the usual exposure duration in video broadcast advertisements. The no exposure condition served as the control. The object shown in the supraliminal and subliminal exposures was a promotional item that we gave to viewers after they watched the video. For the purpose of this study, we selected a promotional item that students were unlikely to have seen prior to the experiment to establish a common baseline of familiarity. We used a Dartman! (see Figure 1). Through pilot tests, we established that the student population was not familiar with the Dartman! but it was an effective promotional item because they enjoyed playing with it. The promotional item was imprinted with the name of the university and College of Business. The top of the Dartman! is removable and conceals a sticky head that, when thrown, sticks to surfaces (e.g., wall, ceiling, computer monitor, refrigerator). It can be used, for example, as a memo holder, a novelty item that sticks to a wall, or an item to place on a bookcase.

We created a fast-paced, 2-min musical promotional video featuring selected alumni and the current year's recipients of outstanding student awards from the College of Business. The video opened with the words *From here* followed by historical and contemporary scenes of the College of Business: students walking to the building, instructors presenting to their classes, and students participating in a sports rally. The word *Anywhere* then followed. Photos of alumni were displayed successively. Words framed the portrait: Across the top was the statement *College of Business Administration*; at the bottom was the alumnus name, position, and company, the left side provided the statement *class of*; and



FIGURE 1. The Dartman! promotional item.

the right side provided the alumnus graduation class. Then the word *Everywhere* was displayed, followed by more portraits of alumni. Last, the words *No Limits* were displayed, and the outstanding student award recipients with captions of their names were successively listed at the end of the promotion. The video ended with a recap of the words *From here...anywhere...everywhere...NO LIMITS College of Business*, [name of the university].

The three versions of the promotional spot were identical and differed only in the inclusion or exclusion and length of exposure time of the promotional item. In the subliminal piece, the promotional item was shown three times, once each as the words *From here, Anywhere, Everywhere* were shown. The promotional item received three exposures of 3 frames (1/10 s) each. In the normal exposure video, the promotional item was shown three times at the same locations as in the subliminal piece but was on the screen each time for 8 frames (approximately 1/4 s). It was also shown at the end of the video, when it flew up from the bottom of the screen to the top of the screen as the words *From here, Anywhere, Everywhere, College of Business Administration*, [name of university] were shown. This Dartman! was on the screen for 9 s (72 frames). The no exposure condition did not show the promotional item.

Participants

Students are especially well positioned to judge the effectiveness of the video to recruit future students, so we recruited students for this study. We randomly assigned students from six sections of introductory marketing classes ($n = 45, 37, 32, 47, 39, 48$; 32% men and 68% women; average age = 23.9 years) to one of the experimental conditions. Because students can enroll in only one section of the introductory marketing class, the likelihood of students participating in more than one experimental condition was eliminated. In addition, discussion among the students about the experiment, their participation, or the promotional item was minimized because the students enrolled in the different sections were not likely to interact with one another. Last, the status of the introductory marketing course as a required course for all business majors ensured that the classes would be representative of the population of business majors at the university. Many of the students are community college transfer students, and the introductory marketing course is among their first business course. Because they have recently joined the campus, they are an appropriate representative sample of the target population for whom this video was made.

Procedure

We informed the participants that the College of Business had developed a promotional video for recruiting. Participants watched the video and evaluated the format on the survey we provided. We then informed participants that, in

addition to showing the video to selected audiences, the College of Business planned to give away a promotional item (the Dartman!), which we then showed the participants and demonstrated its use. In the endowed condition, the researcher gave the participants the dartman to keep. In the unendowed condition, the researcher gave the participants the dartman to examine and then collected it. Participants evaluated the Dartman! on the survey we provided. The survey for both the video and the promotional item was comprised of a series of 5-point Likert-type attitude scales, with responses ranging from 1 (*very effective*) to 5 (*very ineffective*). To provide legitimacy to the cover story, we added six filler items to the survey in which participants evaluated the format, music, theme, length, appearance, and appeal of the video. We also added a question measuring the overall effect of the video. In the survey for the promotional item, participants evaluated the Dartman! on six filler items (uniqueness, memorable, fun, usefulness, appropriateness, and appeal) and on overall effectiveness. To provide a manipulation check on the novelty of the promotional item, participants indicated whether they had ever seen the item before participating in the study. We eliminated the scores of five participants who indicated that they had seen the item or a similar item before the study. Last, participants in the endowed condition indicated the price at which they were willing to sell the Dartman!, and participants in the unendowed condition indicated the price at which they were willing to buy the Dartman! The survey concluded with demographic questions about age, gender, and class standing.

Results

We performed a 2×3 multivariate analysis of variance (MANOVA) using possession (endowed and unendowed) and exposure (no exposure, subliminal, and supraliminal) as the independent variables and answers to survey questions about overall effectiveness of the promotion item and selling or buying price for the promotional item as the dependent variables. We evaluated the multivariate F ratios using the Wilks's lambda criterion. The two-way multivariate interaction was not significant for the rating of the preference or object valuations of the promotional item, $F(4, 464) = .254, ns$. Thus, H_3 and H_4 were not supported.

An analysis of variance (ANOVA) for the endowment condition revealed significant findings for the selling or buying price, $F(1, 232) = 6.536, p = .011, \eta^2 = .0271$. The average selling price (82 cents, $SD = 79.08$) was higher than was the average buying price (52 cents, $SD = 70.47$). These results demonstrate the endowment effect and support H_2 . The ANOVA for the endowment condition was not significant for the rating of object preference, $F(1, 232) = .534, ns$.

The ANOVA for the exposure condition was significant for object preference, $F(1, 233) = 3.653, p = .027, \eta^2 = .051$. Tukey post hoc tests revealed that subliminal exposure ($M = 2.36, SD = .917$) resulted in greater object preference than did supraliminal or normal exposure ($M = 2.54, SD = .925$) and than did

no exposure (control condition; $M = 2.90$, $SD = .982$). Tukey post hoc tests revealed that supraliminal or normal exposure ($M = 2.54$, $SD = .925$) resulted in greater object valuation than did no exposure ($M = 2.90$, $SD = .982$). These results support H_1 . The ANOVA for object valuation did not differ by the exposure condition, $F(1, 232) = .149$, ns .

Discussion

By performing a MANOVA and ANOVA for the endowment condition and an ANOVA for the exposure condition, we found no difference in the average ratings for the advertisements (no exposure = 2.79, $SD = 1.02$; subliminal exposure = 2.80, $SD = 1.08$; supraliminal exposure = 2.81, $SD = 1.09$; endowed = 2.85, $SD = 1.08$; unendowed = 2.80, $SD = 1.03$). The results indicate a close agreement among the groups' rating of the video as neutral (3 on the 5-point Likert-type scale) and suggest that differences in object valuation or object preference ratings were not caused by differential advertisement preferences.

Our results suggest that the mere exposure effect and the endowment effect do not interact to affect object preference. Subliminal exposure produced a stronger mere exposure effect than did supraliminal exposure and no exposure, and supraliminal exposure produced a stronger mere exposure effect than did no exposure. Thus, whereas both subliminal and supraliminal exposures produced a mere exposure effect, the mere exposure effect was stronger in the subliminal condition. One explanation for this finding is that subliminal exposure reflects the stronger effect of the occurrence of a nonconscious process (Bornstein & D'Agostino, 1992). Priming is effective when consumers are not aware that they are being primed (Strahan et al., 2001; Winkielman et al., 2001). Participants' nonconscious, subliminally primed motive was then matched with their subsequent receipt of the promotional item. The object preference likely reflects the match between the prime and the object (Ratneshwar, Mick, & Huffman, 2000; Read & van Leeuwen, 1998). Another explanation for this finding is that the subliminal exposure of the promotional item provided a mere exposure effect while retaining the novelty characteristics of the promotional item (i.e., although consumers were exposed to the Dartman!, their unawareness of the exposure maintained its novelty). In the case of supraliminal exposure, the mere exposure effect was attenuated by the awareness of the novelty characteristics of the promotional item. Thus, the more favorable rating for the Dartman! in the subliminal condition than in the supraliminal condition may have been a result of maintenance of its novelty. The ratings of such an item likely reflect its inherent novelty.

We did not find an interaction effect of the mere exposure effect and the endowment effect on object valuation. However, we found a significant main effect of the endowed condition on object valuation. These results suggest that either object valuation is not affected by the mere exposure effect or, alternatively, these results may be the result of an insufficient manipulation of the mere

exposure effect. It is possible that the mere exposure effect can affect object valuation but that it was not sufficiently strong in this experiment to do so. Future researchers should explore this possibility.

Previous researchers in the mere exposure effect used slides, but we used a video. Video presentation is a realistic representation of advertisements broadcasted on television and the Internet, but it reduces ability to control participants' attention. When slides are used, the experimenter has greater control of participants' attention to the exposed stimuli. In a video, the participant has greater control of attention. Future researchers using videos should weigh the advantages of its realism with the limitations of reduced experimental control of participant attention. Future researchers should compare the effects of video and slide presentations on audience preference, object valuation, attention, and retention. The frequency and duration of exposure we used in this study were within the range used in previous work (Bornstein, 1989). However, future researchers should investigate the effect of different frequencies and durations on consumer behavior. The cover story for our study required the use of an inexpensive novelty item, a Dartman!, as a promotional item. Future researchers should use different objects to determine the generalizability of our findings to items of different monetary values and varying degrees of familiarity.

We demonstrated the impact of the unconscious processes of the mere exposure effect and endowment effect on consumer decision-making behavior. Along with studies within this area of research, our results suggest that the effectiveness of advertisements may be understated if they rely on measures of recall or recognition and point to the need to include instruments that measure the influence of unconscious processes.

REFERENCES

- Bargh, J. A. (1992). Why subliminality does not matter to social psychology: Awareness of the stimulus versus awareness of its influence. In R. F. Bornstein & T. E. Pittman (Eds.), *Perception without awareness* (pp. 236–255). New York: Guilford.
- Bargh, J. A. (2002). Losing conscious: Automatic influences on consumer judgment, behavior, and motivation. *Journal of Consumer Research*, 29, 280–285.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of train construct and stereotype priming on action. *Journal of Personality and Social Psychology*, 71, 230–244.
- Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Troetschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology*, 81, 1014–1027.
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106, 265–289.
- Bornstein, R. F., & D'Agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, 63, 545–552.
- Chartrand, T. L., & Bargh, J. A. (2002). Nonconscious motivations: Their activation, operation, and consequences. In A. Tesser, D. Stapel, & J. Woods (Eds.), *Self-motivation:*

- Emerging psychological perspectives* (pp. 31–41). Washington, DC: American Psychological Association.
- Cohen, G., & Chakravarti, D. (1990). Consumer psychology. *Annual Review of Psychology*, 41, 243–288.
- Fitzsimons, G. J., Hutchinson, J. W., & Williams, P. (2002). Non-conscious influences on consumer choice. *Marketing Letters*, 13, 269–279.
- Fitzsimons, G. J., & Shiv, B. (2001). Nonconscious and contaminative effects of hypothetical questions on decision making. *Journal of Consumer Research*, 28, 224–238.
- Franciosi, R., & Kujal, P. (1996). Experimental tests of the endowment effect. *Journal of Economic Behavior & Organization*, 30, 213–227.
- Jacoby, J., Johar, G. V., & Morrin, M. (1998). Consumer behavior: A quadrennium. *Annual Review of Psychology*, 49, 319–344.
- Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7, 35–51.
- Kahneman, D., Knetsch, J. L., & Thaler, R. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy*, 98, 1325–1348.
- Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, 39, 341–350.
- Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *The American Economic Review*, 79, 1277–1284.
- Knetsch, J. L. (1990). *Derived indifference curves*. Unpublished manuscript, Department of Economics, Simon Fraser University, Burnaby, BC, Canada.
- Ledoux, J. (1996). *The emotional brain*. New York: Simon and Schuster.
- Ratneshwar, S. D., Mick, G., & Huffman, C. (Eds.). (2000). *The why of consumption*. New York: Routledge.
- Read, D., & Van Leeuwen, B. (1998). Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Process*, 76, 189–205.
- Simonson, I., Carmon, Z., Dhar, R., Drolet, A., & Nowlis, S. M. (2001). Consumer research: In search of identity. *Annual Review of Psychology*, 52, 249–275.
- Strahan, E. R., Spencer, S. J., & Zanna, M. P. (2001). *Subliminal priming and persuasion: Striking while the iron is hot*. Unpublished manuscript, Department of Psychology, University of Waterloo, Ontario, Canada.
- Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1, 39–60.
- Tom, G. (2004). The endowment-institutional affinity effect. *The Journal of Psychology*, 136, 160–170.
- Wilson, W. R. (1979). Feeling more than we can know: Exposure effects without learning. *Journal of Personality and Social Psychology*, 37, 811–821.
- Winkielman, P., Berridge, K. C., & Wilbarger, J. (2001). *Doing without feeling: Unconscious affect controls human consumption*. Unpublished manuscript, Department of Psychology, University of Denver, CO.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9, 1–28.
- Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10, 224–228.
- Zola-Morgan, S., Squire, L. R., Alvarez-Royo, P., & Clower, R. P. (1991). Independence of memory functions and emotional behavior. *Hippocampus*, 1, 207–220.

Original manuscript received May 30, 2006
Final version accepted July 24, 2006

CONSUMER PREFERENCE: WANTING & LIKING



A recurring theme in this compendium is whether choices are conscious or unconscious. Both from psychology and neuroscience, we know that we have (at least) two different motivation systems. Some scholars have distinguished between a wanting system and a liking system. The wanting system operates unconsciously, and is known to reflect the operations of a separate, deep brain system, such as the basal ganglia.

Conversely, the *liking system* is related to our overt, conscious hedonic experience, and is reflected in our verbal utterances. This system is thought to be based on the brain's prefrontal regions, such as the orbitofrontal cortex, and possibly the anterior insula.

In most cases, what you want is what you like. But in some instances, those processes break down. Consumer choice is packed with such examples. Think of the choice of eating a chocolate while on a diet. You feel the urge for a snack and the sugar boost, but consciously think it's a bad thing to do. Which of the two will win?

This is a tell tale sign of consumer motivation, and why the whole concept of neuromarketing and consumer neuroscience makes sense. By only asking people, we only have access to the liking system. By exploring the hidden mental value operations of the brain, we can also understand the wanting system.



Investigating the role of the ventromedial prefrontal cortex in the assessment of brands

José Paulo Santos^{1,2*}, Daniela Seixas^{3,4,5}, Sofia Brandão⁶ and Luiz Moutinho⁷

¹ Superior Institute of Maia, Maia, Portugal

² Socius – Research Centre in Economic and Organizational Sociology, Lisbon, Portugal

³ Faculty of Medicine, Department of Experimental Biology, University of Porto, Porto, Portugal

⁴ Instituto de Biologia Molecular e Celular, Porto, Portugal

⁵ Department of Imaging, Centro Hospitalar de Vila Nova de Gaia e Espinho, Vila Nova de Gaia, Portugal

⁶ Magnetic Resonance Imaging Unit, Department of Radiology, São João Hospital, Porto, Portugal

⁷ Department of Management, University of Glasgow, Glasgow, Scotland

Edited by:

Julia Trommershaeuser, New York University, USA

Reviewed by:

Ming Hsu, University of California Berkeley, USA

Jamie D. Roitman, University of California San Francisco, USA

***Correspondence:**

José Paulo Santos, Instituto Superior da Maia, Avenida Carlos Oliveira Campos, Castelo da Maia, 4475-690 Avioso (São Pedro), Maia, Portugal.
e-mail: jpsantos@ismai.pt

The ventromedial prefrontal cortex (vmPFC) is believed to be important in everyday preference judgments, processing emotions during decision-making. However, there is still controversy in the literature regarding the participation of the vmPFC. To further elucidate the contribution of the vmPFC in brand preference, we designed a functional magnetic resonance imaging (fMRI) study where 18 subjects assessed positive, indifferent, and fictitious brands. Also, both the period during and after the decision process were analyzed, hoping to unravel temporally the role of the vmPFC, using modeled and model-free fMRI analysis. Considering together the period before and after decision-making, there was activation of the vmPFC when comparing positive with indifferent or fictitious brands. However, when the decision-making period was separated from the moment after the response, and especially for positive brands, the vmPFC was more active after the choice than during the decision process itself, challenging some of the existing literature. The results of the present study support the notion that the vmPFC may be unimportant in the decision stage of brand preference, questioning theories that postulate that the vmPFC is in the origin of such a choice. Further studies are needed to investigate in detail why the vmPFC seems to be involved in brand preference only after the decision process.

Keywords: neuromarketing, brands, emotion, preference, multivariate analysis, fMRI

INTRODUCTION

In the last few years several articles were published involving a new approach to the study of brands using neuroscientific techniques. One of these first studies used photographs of soft drinks where brands figured explicitly, inducing preference judgments (Paulus and Frank, 2003). These authors hypothesized that a specific area in the prefrontal cortex, the ventromedial prefrontal cortex (vmPFC), was critical for everyday preference judgments. In fact, they found important activations in this brain region when participants selected preferred soft drinks in contrast with a visual discrimination task of the same stimuli (liquids contained in bottles or glasses). Also investigating brands, Deppe et al. (2005), largely based on the work of Damásio, Bechara, and co-workers (Damásio, 1994; Bechara et al., 1997, 1999; Bechara and Damásio, 2005), proposed a dichotomic theory in economic decision-making, “(...) one chain involving emotional experience (...) and another one based on reasoning strategies” (p. 180). Deppe et al. (2005) propose the vmPFC to be central in the processing of emotions during decision-making, whereas brain regions associated with working memory could sustain reasoning.

Koenigs and Tranel (2008) recruited patients with a specific damage of the vmPFC to select soft drinks in two conditions: blinded or brand-cued. While healthy controls and patients with damage in other brain areas changed their soda preference from the blinded drinks to the brand-cued ones, patients with a lesion

in the vmPFC persisted in their original choice, ignoring brand information, showing that the vmPFC is also necessary in the integration of information in the decision-making process.

In addition the vmPFC was found to be important in signaling risk probabilities (Tom et al., 2007; Rangel et al., 2008). Fellows and Farah (2007) again working with patients with vmPFC impairment, suggested that this brain region is necessary for all sorts of choice tasks, either uncertain (including risky or ambiguous situations), or certain.

However, there is still controversy in the literature regarding the function of the vmPFC in decision-making in general, and in brand preference in particular. For example, Schaefer and Rotte (2007) did not report activations in this brain region when sport and luxury car brands (rewarding stimuli) were compared with rational choices of car brands. In another study, using fNIRS to compare luxury and common handbags assessed individually, Lin et al. (2010) suggest that the cognitive subprocesses that underlie the assessment of branded handbags were only important after the choice was made.

To further elucidate the contribution of the vmPFC in brand preference, we designed a functional magnetic resonance imaging (fMRI) study where subjects assessed positive, indifferent and fictitious brands, testing the participation of the vmPFC in the processing of these different hedonic categories of brands. Moreover, both the period during and after the decision process were analyzed, hoping to unravel temporally the role of the vmPFC.

The designation vmPFC is ambiguous in the literature. The present study relies on the probabilistic atlases *Harvard-Oxford Cortical Structural Atlas* and *Harvard-Oxford Subcortical Structural Atlas* provided by the Harvard Centre for Morphometric Analysis¹. We have considered the vmPFC to include the ventral medial frontal pole, frontal medial cortex, ventral paracingulate gyrus, ventral anterior cingulate gyrus, and subcallosal cortex, limited dorsally by the plane $z = +10$, and laterally by the planes $x = \pm 20$ (MNI152 coordinates).

MATERIALS AND METHODS

GENERAL STRUCTURE

To explore the research question, an event-related fMRI experiment was designed. There were four different stimuli categories, plus the interstimuli interval. Each category was composed by 35 slides (6 s each). The interstimuli interval ranged from 4 until 9 s, in 0.5 s steps. The experiment duration was 1200 s, plus 9 s added in the end to ensure that all of the hemodynamic response was included. The sequence was optimized with Optseq2 software (Athinoula A. Martinos Center for Biomedical Imaging, USA)².

Three of the four stimuli were brands' logos grouped in the following categories: positive, indifferent, and fictitious brands. The fourth stimulus was non-emotional words. During the interstimuli interval participants fixated a cross.

BRAND SELECTION

In order to select the logos for the positive and indifferent brand categories, participants completed an electronic survey in which were shown 200 brand logos, that they had to rate using the *pleasure* and *arousal* dimensions of the PAD – pleasure, arousal, dominance scale (Russell and Mehrabian, 1977; Mehrabian and De Wetter, 1987; Mehrabian, 1995), and the SAM – self assessment manikin, explained in detail elsewhere (Morris, 1995; Bradley and Lang, 2007). Self-reporting emotions is a complex task for most individuals, mainly due to the difficulty in verbalizing such inner states (Chamberlain and Broderick, 2007). SAM is a non-verbal pictorial assessment technique designed to represent each dimension of the PAD scale associated with a person's affective reaction to a certain stimuli. *Dominance* was not included in the brand assessment because with static pictures this dimension correlates with *pleasure* (Bradley and Lang, 2007).

After this task, the responses were screened and categorized according to the following criteria: positive brands if the score was ≥ 7 in the *pleasure* dimension, and ≥ 5 in the *arousal* dimension; indifferent brands if the score was ≥ 4 and ≤ 6 in the *pleasure* dimension, and ≤ 5 in the *arousal* dimension. With this procedure 35 positive and 35 indifferent brands were chosen for each participant, and were randomized to enter the fMRI paradigm.

FICTITIOUS LOGOS

The fictitious brands were brands' logos that did not exist in the market. Each logo was designed by a marketer made to resemble a real one, making it plausible for the consumer. The fictitious brands did not represent a particular type of product. Instead, logos with assorted shapes, colors, and fonts suggesting different products and services were used (examples in Figure 1).

¹<http://www.cma.mgh.harvard.edu>

²<http://surfer.nmr.mgh.harvard.edu/optseq/>



FIGURE 1 | Examples of some of the logos used as fictitious stimuli.

NON-EMOTIONAL WORDS

The fourth stimulus (a second baseline) was non-emotional words: determiners, conjunctions, prepositions, or adverbs. Importantly, nouns or verbs that could evoke emotions, objects, or actions were not used. With this stimulus we hoped to avoid meditation during the fMRI task (Gusnard and Raichle, 2001; Beckmann and Smith, 2005; De Luca et al., 2006), that could cloud possible self-reflexive processes elicited by brands (Yoon et al., 2006).

STRUCTURING THE PARADIGM

The structure of the paradigm was the same for all participants. The paradigm sequences were programmed with SuperLab 4.0 software (version 4.0.6b; Cedrus Corporation, USA)³.

INSTRUCTIONS FOR THE SCANNING SESSION

Depending on the type of stimulus visualized, the participants were instructed to either rate hedonically the brand (as positive, negative, indifferent, or unknown), to read covertly non-emotional words, or just to fixate a cross. Participants made their choices using a button box (model Lumina LU400-PAIR; Cedrus Corporation, USA)⁴.

HUMAN SUBJECTS

The participants were 18, 7 healthy male and 11 healthy female volunteers, right handed, with neither history of neurological nor psychiatric disturbances (mean age 28.2 ± 6.9 years, 19–41 years). Informed consent was obtained in all cases. A safety form for magnetic resonance imaging was filled by every participant. After each session the participants were debriefed. This research project was performed according to the Declaration of Helsinki and was approved by the local Ethics Committee.

DATA ACQUISITION

Functional images with axial orientation were obtained using a T2*-weighted EPI sequence in a Siemens® Magnetom Trio high field (3 T) MRI scanner (Siemens AG, Germany; TR = 3000 ms, TE = 30 ms, 64×64 matrix, FOV = 192 mm, 3.0 mm axial slices). The order of acquisition of the slices was interleaved, and they covered the whole brain. The study consisted in one session where 407 volumes were acquired. The first four volumes were discarded to ensure pulses stabilization.

³<http://www.superlab.com>

⁴<http://www.cedrus.com>

A whole brain anatomical structural scan was acquired also for each volunteer, using a T1-weighted MPRAGE protocol (256×256 matrix, $\text{FOV} = 256 \text{ mm}$, 3.0 mm axial slices), for co-registration purposes. Gradient field mapping was additionally acquired for image quality control.

IMAGE ANALYSIS

Functional magnetic resonance imaging data processing was carried out using FEAT (FMRI Expert Analysis Tool) version 5.98, a model-based GLM (general linear model) analysis tool, and also using probabilistic independent component analysis (PICA; Beckmann and Smith, 2004) as implemented in MELODIC (Multivariate Exploratory Linear Decomposition into Independent Components) version 3.09, a model-free analysis tool, both part of FSL – FMRI's Software Library⁵ (Smith et al., 2004; Woolrich et al., 2009).

General linear model analysis – common procedures

In the FEAT analysis, the following pre-statistics processing was applied: motion correction using MCFLIRT (Jenkinson et al., 2002); slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5 mm; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; high pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 30.0 s). Stimuli were convolved with a gamma function with canonical values (phase 0 s, SD 3 s, and mean lag 6 s). To account for variations, temporal derivatives were added for every explanatory variable (EV), in order to achieve a better fit between the signal and the stimuli convolved hemodynamic responses. Time-series statistical analysis was performed using FILM with local autocorrelation correction (Woolrich et al., 2001). Registration to high-resolution structural and/or standard space images was done using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002).

⁵<http://www.fmrib.ox.ac.uk/fsl>

At the individual level two different strategies of analysis were used for comparison. The first strategy was a traditional approach where the hemodynamic response was investigated during the complete time window of the stimulus (6 s). In the second approach the stimulus duration was divided in two: the period before the response (decision-making), and the period after the response (passive period; see Figure 2).

General linear model analysis – conventional stimulus analysis

Before the scanning session, participants assessed a set of 200 brand logos, from which the *positive* and *indifferent* stimuli were extracted. Then, during the scanning, participants rated again the brands. In the first model, in which the whole of the stimulus duration was considered, 13 EVs were included: the three types of stimulus (positive, indifferent, and fictitious logos) times the four possible ratings (positive, indifferent, negative, and unknown), and the non-emotional words.

Most of the assessments were consistent between the two study sessions, before and during the scanning (see Brand Selection), but some of the possible combinations received little or even no ratings. Although all the possibilities were modeled with EVs aiming to explain most of the variance, only those that were consistent between sessions, i.e., positive brands that were rated as positive during the scanning session (PosPos), indifferent brands that were rated as indifferent (IndInd), or fictitious logos that were rated as unknown inside the scanner (NoBUnk) were considered in the analysis. Hence, at the individual level analysis, stimuli, and baseline were compared, resulting in the following contrasts: positive > indifferent, positive > unrecognized logos, and indifferent > unrecognized logos.

General linear model analysis – stimulus detailed analysis

In the second model, 25 EVs were considered: the three types of stimulus (positive, indifferent, and fictitious logos), times the four possible ratings (positive, indifferent, negative, and unknown), times the two epochs (before and after button pressing), and the non-emotional words.

At the individual level and as before, stimuli and baseline were subtracted, resulting in the following six contrasts (ar: after response; br: before response): positive br > indifferent br, positive

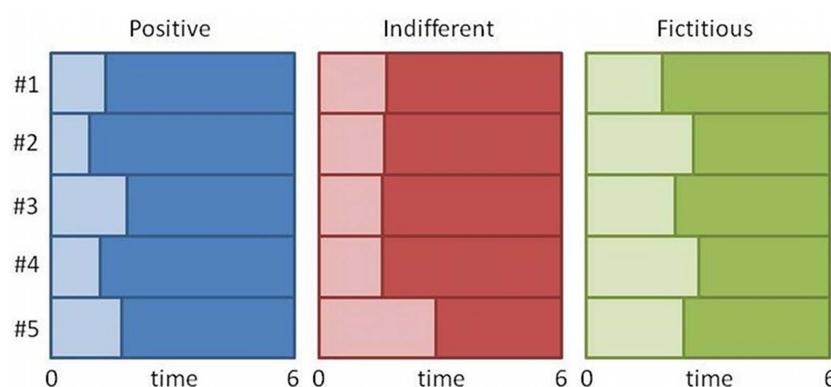


FIGURE 2 | Splitting the duration of the stimulus for one subject. The figure represents the splitting of the first five stimuli of each category (positive, indifferent, and fictitious logos). Lighter areas represent the period until the response (during decision), and darker areas represent the period after the response (passive visualization of the stimulus).

br > unrecognized logos br, indifferent br > unrecognized logos br, positive ar > indifferent ar, positive ar > unrecognized logos ar, and indifferent ar > unrecognized logos ar.

General linear model analysis – group analysis

For both models, group analysis was performed with FLAME (FMRIB's Local Analysis of Mixed Effects) stage 1 and stage 2 with automatic outlier detection (Beckmann et al., 2003; Woolrich et al., 2004; Woolrich, 2008). At this level, group means were calculated from the individual level contrasts.

Z (Gaussianized T/F) statistic images were thresholded using clusters determined by $z > 2.3$ and a (corrected) cluster significance threshold of $p = 1.00$ (Worsley, 2001). Only clusters with more than 50 voxels survived the threshold.

Probabilistic independent component analysis

The following data pre-processing was applied: masking of non-brain voxels, voxel-wise de-meaning of the data, and normalization of the voxel-wise variance. Pre-processed data were whitened and projected into a 164-dimensional subspace using probabilistic Principal Component Analysis where the number of dimensions was estimated using the Laplace approximation to the Bayesian evidence of the model order (Minka, 2000; Beckmann and Smith, 2004). The whitened observations were decomposed into sets of vectors, which describe signal variation across the temporal domain (time-courses), the session/subject domain and across the spatial domain (maps) by optimizing for non-Gaussian spatial source distributions using a fixed-point iteration technique (Hyvärinen, 1999). Estimated component maps were divided by the SD of the residual noise and thresholded by fitting a mixture model to the histogram of intensity values (Beckmann and Smith, 2004).

The EVs basic shapes convolved with a gamma function and including temporal derivatives were concatenated for all the participants in the same order that time-courses were entered in MELODIC, and the same contrasts used in FEAT were computed. The parameter estimates of each spatial independent component (164 total) were then calculated and tested using GLM for each case (see Figure 3), and so the selection of significant spatial independent components was based on statistical criteria.

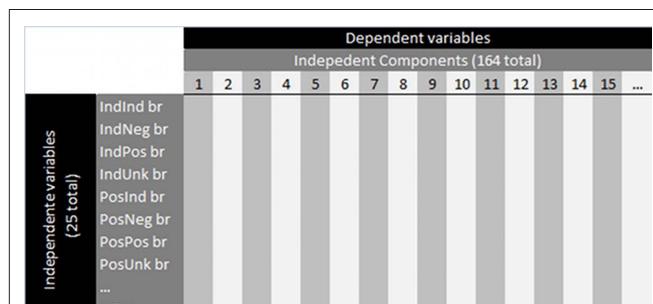


FIGURE 3 | Illustration of the application of a GLM analysis to each of the 164 independent components yielded by MELODIC. For each IC, 25 independent variables were modeled: the three types of stimulus (positive, indifferent, and fictitious logos), times the four possible ratings (positive, indifferent, negative, and unknown), times the two epochs (before and after button pressing), and the non-emotional words.

RESULTS

CONSISTENCY IN THE ASSESSMENTS BETWEEN SESSIONS

Most of the ratings were coherent from one session to the other. Results are summarized in Table 1. Five hundred fifty-four fictitious brands' logos out of 630 (87.9%) were rated as *unknown*, 590 positive brands out of 630 (93.7%) were rated as *positive*, and 427 indifferent brands out of 630 (67.8%) were again rated as *indifferent*.

RESPONSE TIME

The graph in Figure 4 depicts the distribution of the subjects' choices by response time. Response times were shorter with positive ratings (1546 ms) than indifferent (2370 ms) or fictitious ratings (2334 ms), suggesting a delayed decision process with the last two ratings. These differences are significant between positive and indifferent ratings ($F_{426,589,0,01} = 1.702 - p\text{-value} < 0.000001$), and positive and fictitious ratings ($F_{553,589,0,01} = 1.708 - p\text{-value} < 0.000001$), but not significant between indifferent and fictitious ratings ($F_{553,426,0,01} = 1.004 - p\text{-value} = 0.969508$).

GENERAL LINEAR MODEL ANALYSIS

In the conventional GLM analysis (the whole of the stimulus duration), the vmPFC was significantly and extensively activate for the contrasts positive versus indifferent or fictitious logos (Figure 5).

Figure 6 represents the stimulus detailed analysis for the same contrasts. For the period before the response (decision stage) the vmPFC tendentiously deactivated. Conversely, after button pressing, i.e., after the decision was made and while subjects were passively visualizing the stimulus, the vmPFC was active.

Four local maxima from the cluster in the vmPFC in the contrast positive versus indifferent in the conventional analysis were selected for further analysis. The parameter estimates of these voxels are represented in Figure 7 both for the conventional analysis and for the stimulus detailed analysis. For the conventional GLM analysis, all the four local maxima significantly activated when positive brands were involved. On the contrary, in the stimulus detailed analysis there were deactivations, more prominent in the anterior subregions (ventral paracingulate gyrus and ventral medial frontal pole). After the response, however, the vmPFC was extensively activate.

PROBABILISTIC INDEPENDENT COMPONENT ANALYSIS

The 164 ICs yielded by PICA account for 86.95% of the variability.

To select the relevant ICs the criteria were: the z statistics of the contrast between the parameter estimates of the positive brands versus the parameter estimates of the indifferent brands, the

Table 1 | Assessments made during the scanning sessions separated according to the type of stimuli.

Stimuli	Recorded ratings					Total
	Positive	Indifferent	Negative	Unknown	No answer	
Positive	590	29	3	6	2	630
Indifferent	82	427	74	44	3	630
Fictitious	33	36	2	554	5	630
Total	705	492	79	604	10	1890

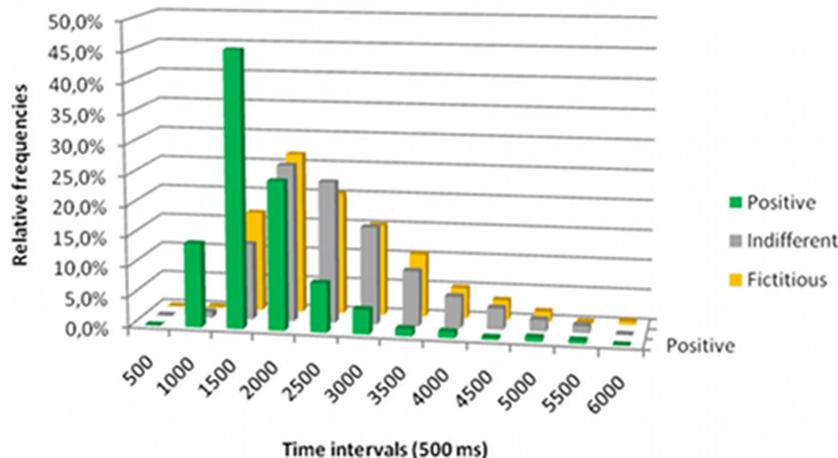


FIGURE 4 | Relative frequency of response times obtained during the scanning session grouped in 500 ms intervals.

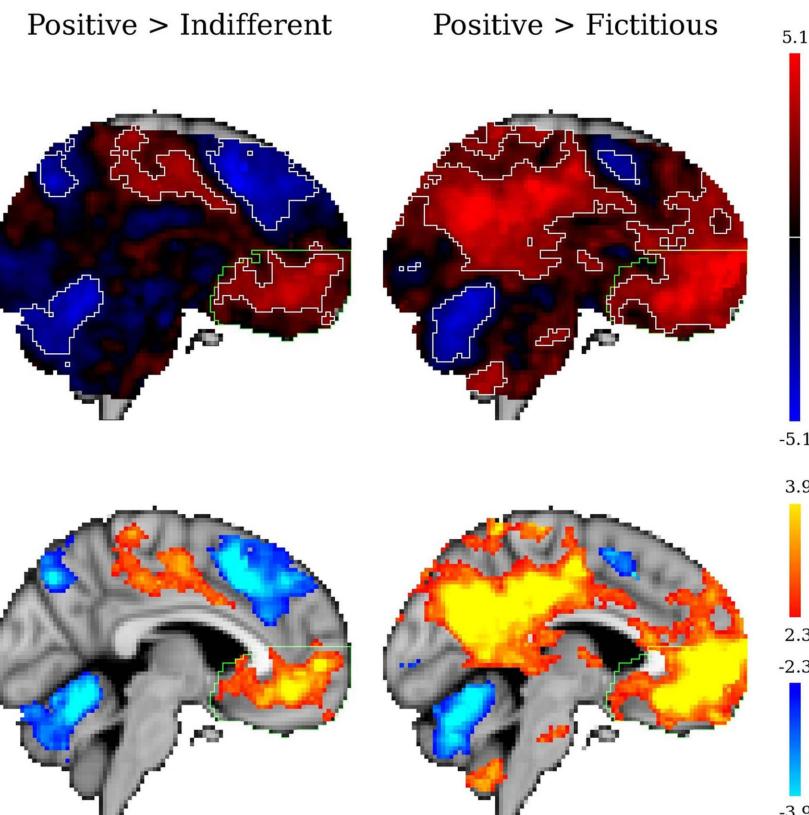
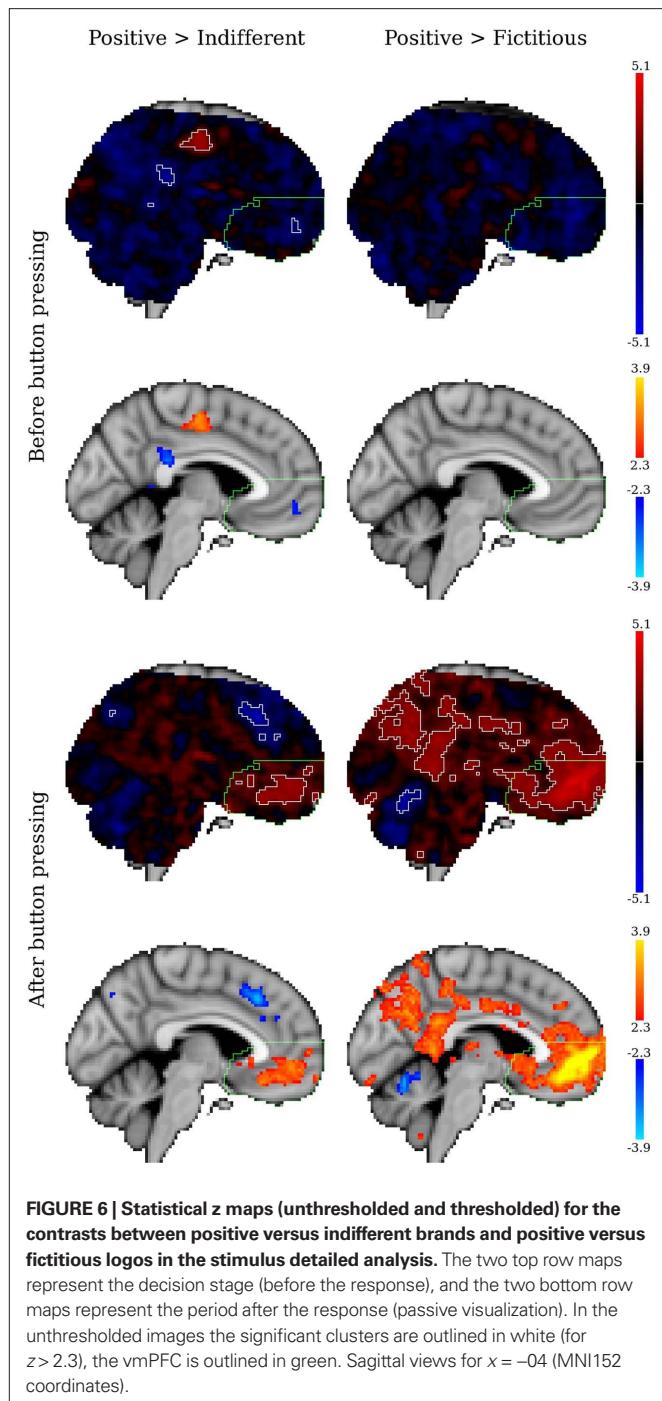


FIGURE 5 | Statistical z maps (unthresholded in the upper row and thresholded in the lower row) for the contrasts positive versus indifferent brands and positive versus fictitious logos in a conventional GLM analysis. In the unthresholded images the significant clusters are outlined in white (for $z > 2.3$), and the vmPFC is outlined in green. Sagittal views for $x = -04$ (MNI152 coordinates).

fictitious logos, and the non-emotional words had to be superior to 2.3 in all the three cases, or inferior to -2.3 in all three cases. This procedure was implemented in the two situations, before and after the response. In this way it was guaranteed that the ICs selected would be significantly more active or more deactivated for positive brands than in the remaining cases.

Neither before the response nor after the response there were ICs with z value inferior to -2.3 . On the contrary, two ICs (#17 and #152) had all the considered z values superior to 2.3 in the situation before the response, and four other (#24, #49, #96, and #135) had z values superior to 2.3 in the situation after the response. Only IC #24 included brain activations in the vmPFC. The z values for the



three cases are reported in **Table 2**. Three slices of IC #24 are represented in **Figure 8**. Besides the vmPFC activation this network also includes active voxels in the precuneus, posterior cingulate gyrus, right and left anterior divisions of the middle temporal gyrus, and deactivation in the occipital fusiform gyrus.

Table 2 also reports the z values for all the ICs that encompass at least one of the local maxima voxels considered in **Figure 7** (activated or deactivated). Only IC #24 has this statistic consistently and significantly positive (for the situation after response). However,

ICs #15 and #22 were significantly positive for the situation after the response in the contrast with other logos (indifferent or fictitious). In the situation before the response, IC #132 had significantly negative z values for the contrasts with logos.

DISCUSSION

Most of the neuroimaging studies involving brands use paradigms consisting of choices between pairs of brands or products, i.e., both stimuli are presented simultaneously and subjects have to choose one or the other. However, the structure adopted in our study is different, we believe closer to everyday life; each brand is presented one at a time, meaning that subjects decide about the hedonic value of a particular brand not by comparison. For example, when a consumer chooses a product from a supermarket shelf, s/he does not collect first all the available items and then choose. On the contrary, there is a previous intention summarized in a concept named consideration set, or evoked set (Roberts and Lattin, 1991; Shocker et al., 1991; Petrof and Daghfous, 1996). The consumer confronts each possibility in the shelf against the consideration set until one brand/product is preferred. Thus, the process is not a simple choice among several options, but instead an assessment of the fit between one option and the inner expectations that were previously constructed.

Damásio (1994) from his observations in neurologically impaired patients, proposed that the prefrontal cortex is a crucial structure in decision-making; the vmPFC in particular is thought to be important in decisions of preference including preference for certain brands (Paulus and Frank, 2003; Deppe et al., 2005; Knutson et al., 2007; Koenigs and Tranel, 2008; Luu and Chau, 2009). The results of our conventional GLM analysis, which included data acquired both before and after decision of brand preference, corroborate these findings: activation of the vmPFC was found when comparing positive with indifferent or fictitious brands. However, when we dissected the subjects' responses and isolated the decision-making period from the moment after the response, we found that, especially for positive brands, the vmPFC was more active after the choice than during the decision process itself, challenging some of the existing literature. And this result was supported both by the GLM time-split analysis and by the PICA analysis.

During the decision process itself, i.e., before the response, the vmPFC was less active for positive brands than for indifferent or fictitious logos. Conversely, the vmPFC was more active after the brand choice was made. Considering the four local maxima in the vmPFC (the subcallosal cortex, the frontal medial cortex, the ventral paracingulate gyrus, and the ventral medial frontal pole), although they were also involved in the conventional analysis when it corresponded to all the period when the stimulus was present, the same voxels of the vmPFC were deactive during the decision period until the response, but active after the response. This pattern was not found with indifferent brands (that subjects recognized as having some meaning to them, but that were not preferred), with fictitious logos (visualized for the first time and about which, likewise, subjects could not have a preformed opinion), and also with non-emotional words.

One of the ICs obtained with the multivariate model-free analysis (PICA) was significantly more relevant in the choice of positive brands than indifferent brands, fictitious logos, or non-emotional

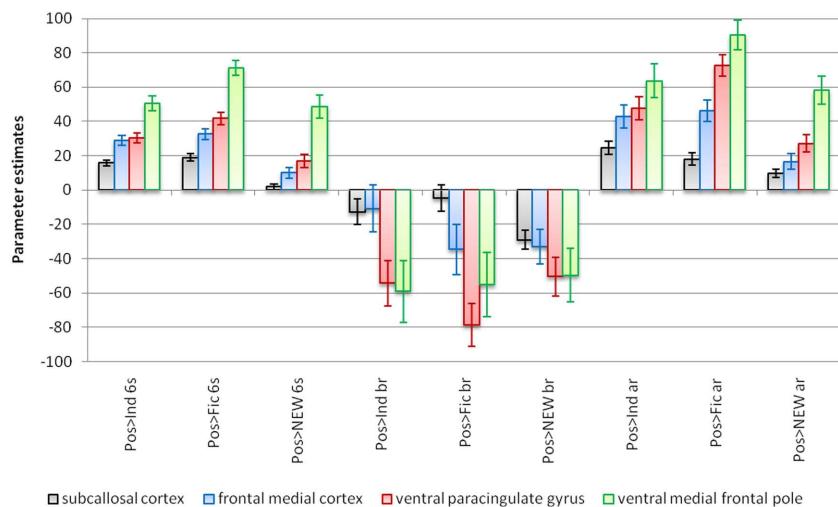


FIGURE 7 | Parameter estimates for the stimuli in four local maxima in the vmPFC (subcallosal cortex: -6, 32, -10; frontal medial cortex: 2, 36, -14; ventral paracingulate gyrus: -2, 48, -2; ventral medial frontal pole: -2, 58, 4). The bar graphs identified with the suffix 6 s are the conventional GLM-based analysis of fMRI data. The bar graphs identified with the suffix br refer to the

participation of the voxel before the response (decision stage). The bar graphs identified with the suffix ar refer to the participation after the decision instant but before the stimulus offset. Pos: positive; Ind: indifferent; Fic: fictitious; NEW: non-emotional words (baseline). MNI152 coordinates. Error bars correspond to confidence intervals of 95%.

Table 2 | Statistic z for all the ICs that had at least one voxel activated or deactivated among those considered in Figure 7.

	IC								
	15	22	24	41	50	89	104	110	132
Pos > Ind br	-0.967	0.227	-1.876	-1.940	-1.588	0.476	-1.581	-0.239	-3.143
Pos > Fic br	-2.560	0.329	1.441	2.296	0.471	0.573	-0.463	3.269	-2.497
Pos > NEW br	-7.146	-1.021	0.275	2.961	1.417	0.388	0.413	-0.358	-1.753
Pos > Ind ar	4.805	3.136	2.562	-2.241	0.819	3.103	-0.173	1.348	-1.353
Pos > Fic ar	4.423	5.432	5.169	2.282	0.520	1.389	2.588	1.487	-0.448
Pos > NEW ar	-4.001	0.693	2.892	-1.790	-1.562	-3.278	-4.711	3.839	-3.542

Pos, positive; Ind, indifferent; Fic, fictitious; NEW, non-emotional words (baseline); br refers to the participation of the voxel before the response (decision stage); ar refers to the participation after the decision instant but before the stimulus offset.

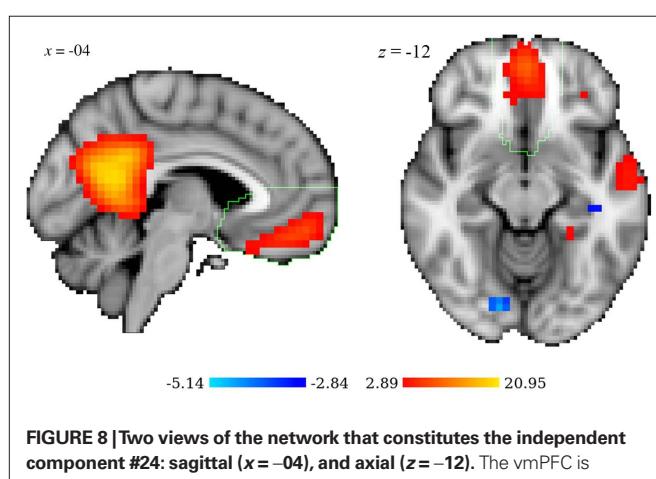


FIGURE 8 | Two views of the network that constitutes the independent component #24: sagittal ($x = -04$), and axial ($z = -12$). The vmPFC is outlined in green. Radiological convention. MNI152 coordinates.

words. IC 24 showed extensive activations in the vmPFC, among other brain structures (Figure 8). This network was significantly more active with preferred brands than with indifferent brands, fictitious logos, or non-emotional words only after the response, which reinforces the fact that although important in decisions of preference, the vmPFC is only so after the decision-making process itself. The analysis of the participation of the vmPFC in brain networks represented in other ICs corroborates this hypothesis, because none of the ICs had consistent or significant statistics to support the participation of the vmPFC in the period before the response.

The results of the present study seem to contradict some of the existing theories on the role of the vmPFC in the decision process. On the other hand, our data are supported by Lin et al., (2010) work in which the brand stimuli were also presented one at a time, suggesting as well a late participation of the vmPFC in preference decision-making; or by Li et al.

(2010) study that used fMRI and the Iowa Gambling Task to investigate the neural correlates of decision-making. They have demonstrated a group of brain regions that included the dorsolateral prefrontal cortex for working memory, and the insula and posterior cingulate cortex for representations of emotional states. However, the vmPFC was not part neither of the memory nor the emotional networks, but instead was coupling the two processes.

REFERENCES

- Bechara, A., and Damásio, A. R. (2005). The somatic marker hypothesis: a neural theory of economic decision. *Games Econ. Behav.* 52, 336–372.
- Bechara, A., Damásio, H., Damásio, A. R., and Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481.
- Bechara, A., Damásio, H., Tranel, D., and Damásio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.
- Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *Neuroimage* 20, 1052–1063.
- Beckmann, C. F., and Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans. Med. Imaging* 23, 137–152.
- Beckmann, C. F., and Smith, S. M. (2005). Tensorial extensions of independent component analysis for multisubject FMRI analysis. *Neuroimage* 25, 294–311.
- Bradley, M. M., and Lang, P. J. (2007). "The international affective picture system (IAPS) in the study of emotion and attention," in *Handbook of Emotion Elicitation and Assessment*, eds J. A. Coan and J. J. B. Allen (New York: Oxford University Press), 29–46.
- Chamberlain, L., and Broderick, A. (2007). The application of physiological observation methods to emotion research. *Qual. Mark. Res. Int. J.* 10, 199–216.
- Damásio, A. R. (1994). *Descarte's Error – Emotion, Reason and the Human Brain*. New York: Penguin Putnam.
- De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., and Smith, S. M. (2006). FMRI resting state networks define distinct modes of long-distance interactions in the human brain. *Neuroimage* 29, 1359–1367.
- Deppe, M., Schwindt, W., Kugel, H., Plassmann, H., and Kenning, P. (2005). Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J. Neuroimaging* 15, 171–182.
- Fellows, L. K., and Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb. Cortex* 17, 2669–2674.
- Gusnard, D. A., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nat. Rev. Neurosci.* 2, 685–694.
- Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* 10, 626–634.
- Jenkinson, M., Bannister, P. R., Brady, J. M., and Smith, S. M. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841.
- Jenkinson, M., and Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156.
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., and Loewenstein, G. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.
- Koenigs, M., and Tranel, D. (2008). Prefrontal cortex damage abolishes brand-cued changes in cola preference. *Soc. Cogn. Affect. Neurosci.* 3, 1–6.
- Li, X., Lu, Z.-L., D'argembeau, A., Ng, M., and Bechara, A. (2010). The iowa gambling task in fMRI images. *Hum. Brain Mapp.* 31, 410–423.
- Lin, C.-H., Tuan, H.-P., and Chiu, Y.-C. (2010). Medial frontal activity in brand-loyal consumers: a behavior and near-infrared ray study. *J. Neurosci. Psychol. Econ.* 3, 59–73.
- Luu, S., and Chau, T. (2009). Decoding subjective preference from single-trial near-infrared spectroscopy signals. *J. Neural Eng.* 6, 016003.
- Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. *Genet. Soc. Gen. Psychol. Monogr.* 121, 339–361.
- Mehrabian, A., and De Wetter, R. (1987). Experimental test of an emotion-based approach to fitting brand names to products. *J. Appl. Psychol.* 72, 125–130.
- Minka, T. P. (2000). *Automatic Choice of Dimensionality for PCA*. Technical Report 514, Cambridge: MIT Media Lab Vision and Modeling Group.
- Morris, J. D. (1995). Observations: SAM: the self-assessment manikin – an efficient cross-cultural measurement of emotional response. *J. Advert. Res.* 35, 63–68.
- Paulus, M. P., and Frank, L. R. (2003). Ventromedial prefrontal cortex activation is critical for preference judgments. *Neuroreport* 14, 1311–1315.
- Petrof, J. V., and Daghfous, N. (1996). Evoked set: myth or reality? *Bus. Horiz.* 39, 72–77.
- Rangel, A., Camerer, C. F., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556.
- Roberts, J., and Lattin, J. (1991). Development and testing of a model of consideration set composition. *J. Mark. Res.* 28, 429–440.
- Russell, J. A., and Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *J. Res. Pers.* 11, 273–294.
- Schaefer, M., and Rotte, M. (2007). Favorite brands as cultural objects modulate reward circuit. *Neuroreport* 18, 141–145.
- Shocker, A., Ben-Akiva, M., Bocvara, B., and Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: issues, models, and suggestions. *Mark. Lett.* 2, 181–197.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobniak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl. 1), S208–S219.
- Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.
- Woolrich, M. W. (2008). Robust group analysis using outlier inference. *Neuroimage* 41, 286–301.
- Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *Neuroimage* 21, 1732–1747.
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T. E., Beckmann, C. F., Jenkinson, M., and Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45, S173–S186.
- Woolrich, M. W., Ripley, B. D., Brady, J. M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* 14, 1370–1386.
- Worsley, K. J. (2001). "Statistical analysis of activation images," in *Functional MRI: An Introduction to Methods*, eds P. Jezzard, P. M. Matthews, and S. M. Smith (New York, NY: Oxford University Press), 251–270.
- Yoon, C., Gutches, A. H., Feinberg, F., and Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *J. Consum. Res.* 33, 31–40.

In summary, the results of the present study converge in supporting the notion that the vmPFC may be unimportant in the decision stage concerning brand preference, questioning theories that postulate that the vmPFC is in the origin of brand choice. To complement our findings, further studies that challenge as well conventional research design and neuroimaging methodologies are need to investigate in detail why the vmPFC seems to be involved in brand preference only after the decision process.



A neural predictor of cultural popularity

Gregory S. Berns ^{*}, Sara E. Moore

Economics Department and Center for Neuropolicy, Emory University, Atlanta, GA 30322, USA

Received 7 February 2011; accepted 5 May 2011

Abstract

We use neuroimaging to predict cultural popularity — something that is popular in the broadest sense and appeals to a large number of individuals. Neuroeconomic research suggests that activity in reward-related regions of the brain, notably the orbitofrontal cortex and ventral striatum, is predictive of future purchasing decisions, but it is unknown whether the neural signals of a small group of individuals are predictive of the purchasing decisions of the population at large. For neuroimaging to be useful as a measure of widespread popularity, these neural responses would have to generalize to a much larger population that is not the direct subject of the brain imaging itself. Here, we test the possibility of using functional magnetic resonance imaging (fMRI) to predict the relative popularity of a common good: music. We used fMRI to measure the brain responses of a relatively small group of adolescents while listening to songs of largely unknown artists. As a measure of popularity, the sales of these songs were totaled for the three years following scanning, and brain responses were then correlated with these “future” earnings. Although subjective likability of the songs was not predictive of sales, activity within the ventral striatum was significantly correlated with the number of units sold. These results suggest that the neural responses to goods are not only predictive of purchase decisions for those individuals actually scanned, but such responses generalize to the population at large and may be used to predict cultural popularity.

© 2011 Published by Elsevier Inc. on behalf of Society for Consumer Psychology.

Keywords: fMRI; Neuroeconomics; Neuromarketing; Music; Prediction

Introduction

How can we predict popularity? Although superficially a trivial question, the desire for popularity consumes a great portion of the lives of many youths and adults. More than the superficial teenager's quest for popularity, being popular is a marker for social status. Consequently, popularity would seem to confer a reproductive advantage in the evolution of the human species, thus explaining its importance to humans. Such importance extends to economic success as well because goods and services that are popular command higher prices. Although there are good economic and evolutionary rationales for pursuing popularity, predicting who or what becomes popular is a challenging problem. Even so, the ability to predict

popularity is a valuable skill that also can translate into economic success.

In the domain of economic goods, traditional approaches to forecasting popularity rely on standard marketing techniques. These include focus groups, questionnaires, simulated choice tests, and market tests. More recently, however, the widespread use of neuroimaging has raised the possibility of using functional magnetic resonance imaging (fMRI) in the marketing process (Ariely & Berns, 2010). Neuroeconomic research suggests that activity in reward-related regions of the brain, notably the orbitofrontal cortex and ventral striatum is predictive of future purchasing decisions of the individuals who are scanned (Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008; Knutson, Rick, Wimmer, Prelec, & Loewenstein, 2007; Plassmann, O'Doherty, & Rangel, 2007; Plassmann, O'Doherty, Shiv, & Rangel, 2008). For neuroimaging to be useful in either a marketing or branding application, however, these neural signals would need to generalize to a larger group of individuals who themselves were not the direct object of

* Corresponding author at: 1602 Fishburne Dr., Emory University, Atlanta, GA 30322, USA.

E-mail address: gberns@emory.edu (G.S. Berns).

brain scanning. Currently, it is unknown whether the neural signals of a small group of individuals are predictive of the purchasing decisions of the population at large.

Neuroimaging is often touted as a hot new tool for branding (Lindstrom, 2008). Although branding and advertising have been considered in a few neuroimaging papers (Kenning & Plassmann, 2008; Lee, Broderick, & Chamberlain, 2007; Yoon, Gutchess, Feinberg, & Polk, 2006), it is still unknown whether neuroimaging can prospectively reveal whether a particular ad or brand campaign will be effective. In a well-known Coke–Pepsi study, participants who described themselves as Coke-drinkers showed significant activation in the hippocampus and right DLPFC when they were cued about the upcoming drink of Coke (McClure et al., 2004). Self-described Pepsi-drinkers did not have this response. In the absence of brand information, there was no significant difference in preference during a test-taste. This study suggested that any differences in the neural response to the two brands must be culturally derived. Although these results demonstrate that branding does affect brain responses to nominally similar goods, the question of whether brand effectiveness can be predicted in advance remains an open question.

To demonstrate the efficacy of an fMRI study for branding, three conditions must be met. First, the study participants—i.e. the cohort of individuals who are actually scanned—should be representative of the population that is the target of a brand campaign. Second, to truly test whether the neural signals are predictive of brand effectiveness, the scanning must be done before the campaign is launched. Third, metrics of brand effectiveness must be readily available for the target population. For example, these might include sales data, web page views, downloads, internet searches, etc. Finally, although not strictly a condition, it is an open question as to what should actually be scanned during fMRI. If the product can be consumed in the scanner, then the product itself becomes the target. Alternatively, an ad or branding campaign might be presented in the scanner, in which case an abstract association between an ad and a product becomes the scanned target.

One product that meets these requirements is music. Everyone has musical preferences, and most people spend money on this product. Thus, it is straightforward to find people to scan who are representative of the music-consuming public, which is almost everyone. Second, the rise of sites like myspace.com has created a large repository of music which is largely unadvertised and unbranded. Because much of this music is provided directly by the artist, it can be used well in advance of any ad campaign; moreover, the band is the brand. Third, metrics of music success are simple and straightforward: downloads, sales, and ticket receipts. Finally, music is ideally suited to scanning because the act of listening to it is the same as consuming it. Thus, imaging the neural response to music is a direct measure of the consumption experience. Subsequent success is then a combination of quality, branding, and marketing.

In a previous study of adolescents, we measured the interaction of social influence in the form of popularity ratings with the consumption experience of music (Berns, Capra,

Moore, & Noussair, 2010). Using fMRI, we found that although an individual's musical preferences were strongly correlated with activity in the caudate nucleus, the effect of social information varied between participants. The tendency to change one's evaluation of a song was positively correlated with activation in the anterior insula and anterior cingulate, two regions that are associated with physiological arousal and negative affective states. While this earlier study examined the effect of popularity information on individual preferences, here we report a longitudinal analysis in which we examine the relationship between brain responses and popularity of music from the other direction: do neural responses to music in an fMRI study predict subsequent commercial success of the song and artist?

Material and methods

A total of 32 adolescent participants were studied. Five were excluded from the fMRI analyses due to either excessive movement or susceptibility artifacts. Although this was a relatively high exclusion rate compared to adult studies, it was comparable to previous fMRI studies in children and adolescents, who tend to move more than adults (Galvan et al., 2006). Prior to the experiment, they were screened for the presence of medical and psychiatric diagnoses, and none were taking medications. There were 14 female and 13 male participants between the ages of 12 and 17.9 (mean 14.6). Fifteen were Caucasian, eight were African-American, one was Hispanic, and three were "Other." The primary stimuli used were 15-s clips from songs downloaded from MySpace.com. Songs were downloaded between October 23 and November 8, 2006. In order to minimize the possibility that participants would recognize the artists, songs from unsigned musicians or relatively unknown artists were used. A total of 20 songs were downloaded in each of the following genres: Rock, Country, Alternative/Emo/Indie, Hip-Hop/Rap, Jazz/Blues, and Metal (identified by the MySpace category). At the time of download, the number of times each song had been played was recorded, and this was used to calculate the popularity of each song among MySpace users. Each song was converted from MP3 to WAV format and a 15-s clip was extracted that included either the hook or chorus of the song. These 15-s clips were subsequently used in the experiment.

At the beginning of each session, individuals' rankings of musical genres were elicited. Participants were provided with a list of the six musical genres, and were instructed to rank the genres from 1 ("the type you like the best") to 6 ("the type you like the least"). Each participant's top three genres were subsequently used in the experiment. Emory University's Institutional Review Board approved all procedures. Individuals then entered the scanner, and the total scan time was approximately one hour. The scanning was performed on a Siemens 3T Trio. Each subject received a T1-weighted structural image (TR=2600 ms, TE=3.93 ms, flip angle=8, 224×256 matrix, 176 sagittal slices, 1 mm cubic voxel size), a DTI scan (TR=6500 ms, TE=90 ms, flip angle=90, FOV=220 mm, 128×128 matrix, 34 axial slices, 1.7×1.7×2.5 mm voxel size,

6 sets of 12 directional $b=1000$ and 1 $b=0$ images), and 3 functional runs of BOLD-weighting (TR=2000 ms, TE=31 ms, flip angle=90°, FOV=192 mm, 64×64 matrix, 28 axial slices, 3 mm cubic voxel size). Each individual participated in 60 trials. Each trial was divided into two stages; in each stage the subject listened to the same 15-s song clip (Fig. 1a). During stage 1, no popularity information was shown. After listening, subjects were required to rate the song based on (a) how familiar it was and (b) how much they liked it. Both ratings used a 1–5 star scaling system. To prevent the subject from passively accepting a default rating, each rating screen began with 0 stars, which could not be accepted as a final selection. After the rating was entered, stage 2 of the trial took place. The clip was played again, after which the subject provided another likability rating. Twenty songs in

each of the subject's top-three genres were presented in random order throughout the experiment. In 2/3 of the trials, during the second listen, the song's popularity was displayed in the 1–5 star scaling system. The 40 trials in which the popularity display appeared were sequenced randomly among the 60 trials. Only brain activation data from the first listening period was used in the subsequent analysis. As an incentive to accurately reveal their song preferences, each subject received a CD with their top-rated songs.

Nielsen SoundScan was used as the source of post factum popularity information over the three years since the songs were originally chosen for the study. The SoundScan database was searched for information on the performers of each of the 120 songs in the study. Sales data were available for 87 songs,

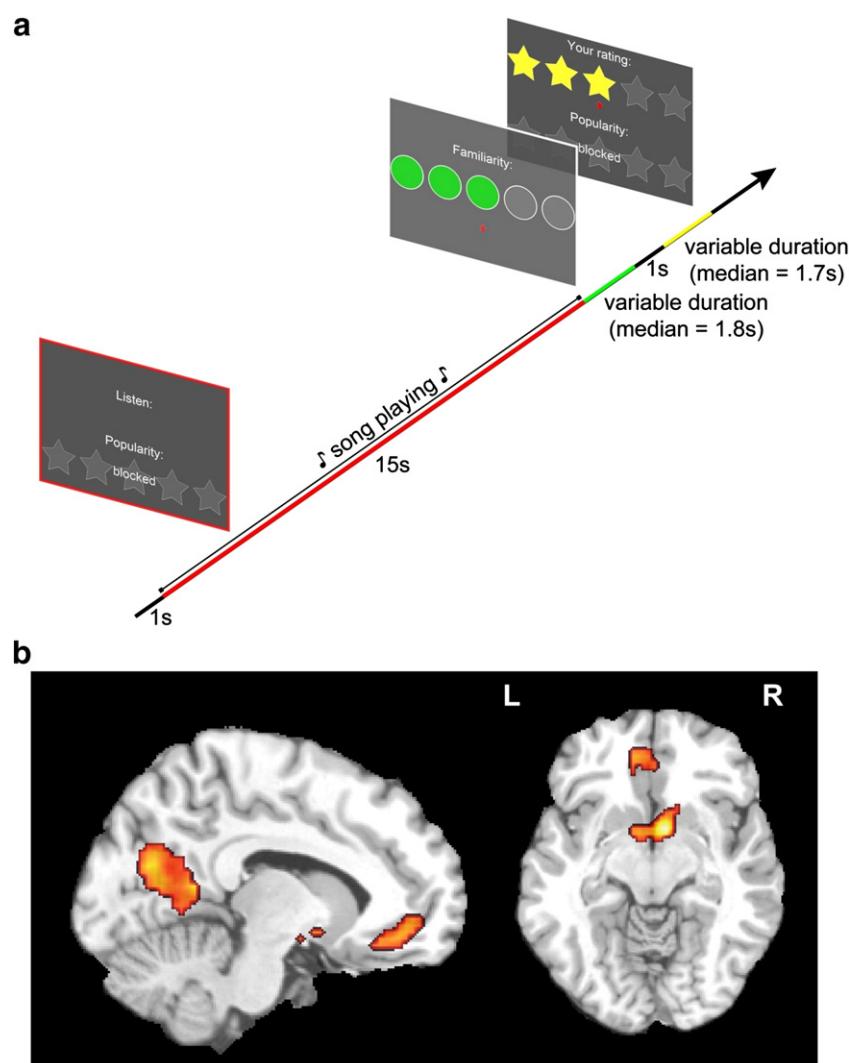


Fig. 1. Design of popular music experiment and brain regions correlated with the average likability of each song. a) Timing of events for a typical trial. Each song was played for 15 s. Each participant heard 60 songs from their favorite 3 genres. Following the song, participants rated the song for familiarity and likability. The trial was then repeated with the average popularity shown on 2/3 of the trials and blocked for 1/3 of the trials. Only the initial listening period (red) was used for subsequent analyses. The first listening period for each of the songs was modeled separately, for a total of sixty 15 second variable duration events for each participant. Second-level models for each of the 120 songs were constructed as one-sample t-tests from the contrast images of the first listening period from the first-level model. A third-level model was then built using the positive contrast images from the second-level model. This model also included a covariate of the first likability rating for each song, averaged over the participants who heard that song. b) Brain regions positively correlated with the average likability of the song from the third-level model ($p < 0.005$, cluster extent ≥ 56 , yielding whole-brain FDR < 0.05) were limited to three areas: cuneus, orbitofrontal cortex, and ventral striatum. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and this data was extracted from the SoundScan database during May 2010. The obtained metric aggregated all recorded sales of the song from release through May 2010, including singles, albums, and compilations.

Preprocessing of the fMRI data was executed in SPM5 (Functional Imaging Laboratory, UCL, London). The preprocessing pipeline consisted of slice timing correction, motion correction, spatial normalization, and smoothing (with an 8 mm Gaussian kernel). A first-level GLM was constructed in SPM5 for each of the 27 participants. The first listening period for each of the 60 songs was modeled separately, for a total of sixty 15-second variable duration events. All of the second listening periods, also 15 s variable duration events, were collapsed into a single condition for each run. All three variable duration rating phases of the trial (familiarity, first likability, and second likability) were also collapsed into one condition to model the act of rating including the button presses. The motion parameters were also included in the model as an effect of non-interest. Second-level models for each of the 120 songs were constructed as one-sample t-tests in SPM5 using contrast images of the first listening period from the first-level model above. Since every participant did not hear every song, the number of contrast images in each of these second-level models ranged from 3 to 23. A third-level model, also a one-sample t-test, was then built in SPM5 using the positive contrast images from the second-level model above. This model also included a covariate of the first likability rating for each song, averaged over the participants who heard that song. We used the likability covariate to identify ROIs. Statistical thresholds were determined based on the estimated smoothness of the 3rd-level contrasts. Using the AlphaSim routine in AFNI, we estimated the combination of height and extent thresholds that yielded a whole-brain FDR<0.05 (10,000 iterations). First, white matter and CSF were masked out using the SPM probabilistic gray matter map. With a gray matter probability>0.6, this results in a mask that retains most gray matter while effectively eliminating most white matter and CSF, which would otherwise inflate the required cluster size. Second, we used 3dFWHMx to estimate the image smoothness from the square root of the masked SPM-generated ResMS image and input into AlphaSim. Finally, using a voxel level threshold of p<0.005, the extent threshold that yielded a cluster level alpha of 0.05 was determined to be k≥56 (Logan & Rowe, 2004; Zhang, Nichols, & Johnson, 2009). The 60% gray matter mask was applied to all contrasts before using these thresholds. Only three brain regions showed a significant correlation between activation and average song likability: cuneus, orbitofrontal cortex (OFC) and ventral striatum/nucleus accumbens (NACC) (Fig. 1b). The activations for each song were then extracted from these regions of interest (ROIs) for subsequent analyses with song sales data. This approach ensures that the ROIs are defined independently from the variable of interest (song sales).

Results

The vast majority of songs in our sample were not commercially successful. The distribution of sales exhibited a long

tail (Fig. 2a), with only three songs' albums meeting the industry standard for "gold" (500,000 units). Given the large number of songs released annually, this is not surprising. To normalize the distribution, sales data were log-transformed for subsequent analyses. First, we checked if either of the subjective song ratings were predictive of future sales, but neither of the two average ratings obtained for each song was correlated with sales data (likability: R=0.110, p=0.313; familiarity: R=0.106, p=0.330), nor was the average genre ranking of each song (p=0.102). This indicates that simple subjective reports collected from study participants may not be good predictors of commercial success.

Although subjective ratings of songs did not correlate with future sales, the activation within the NACC did (Fig. 2b). Log (sales+1) was significantly correlated with the average activation in NACC during the 15-sec listening period (R=0.32, p=0.004). To see which part of the 15-sec period was responsible for this correlation (e.g. initial or final reactions), we tested an alternative model with the listening period divided into three 5-sec segments. None of these three segments exhibited a greater correlation to sales than the whole 15-sec period. This indicates that the mechanism driving the correlation between NACC activity and sales was integrated over the entire listening period.

To further understand the interrelationship between song likability, NACC activity, and sales, we constructed a structural equation model (SEM) (Fig. 3a). The SEM was based on known anatomical connections between the OFC and NACC and their relationships to subjective likability and purchase decisions (Chib, Rangel, Shimojo, & O'Doherty, 2009; Knutson, et al., 2007; Montague & Berns, 2002; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Plassmann et al., 2007; Rolls, 2000). Consistent with this literature, the average song likability had significant path coefficients to both the OFC and NACC, and because of the direct connection between OFC and NACC, this path was also significant. However, the final pathway linking the brain to album sales was mediated only through the NACC. When these relationships were visualized, it became clear that "hit" songs did not result from a specific combination of NACC and OFC activity, but that "non-hits" were associated with a combination of both low OFC and low NACC activity (Fig. 3b). This relationship was quantified through logistic regressions on different hit/non-hit thresholds of sales (Fig. 3c). With thresholds in the range of 15,000 to 35,000 units sold, the logistic model achieved reasonable accuracy in correctly classifying hits and non-hits. For example, with a hit-threshold of 15,000 units, the logistic model correctly classified 80% of the non-hits; however, this came at a cost of missing true hits (but still correctly classified 30% of the hits). To test the possibility that consistency of brain responses might also be predictive of sales, we formulated a model that included a term for the reciprocal of the variance of the brain response across subjects that heard each song; however, this term was not significant [F(1,83)=0.28, p=0.597].

Discussion

Our results demonstrate that not only are signals in reward-related regions of the human brain predictive of individual

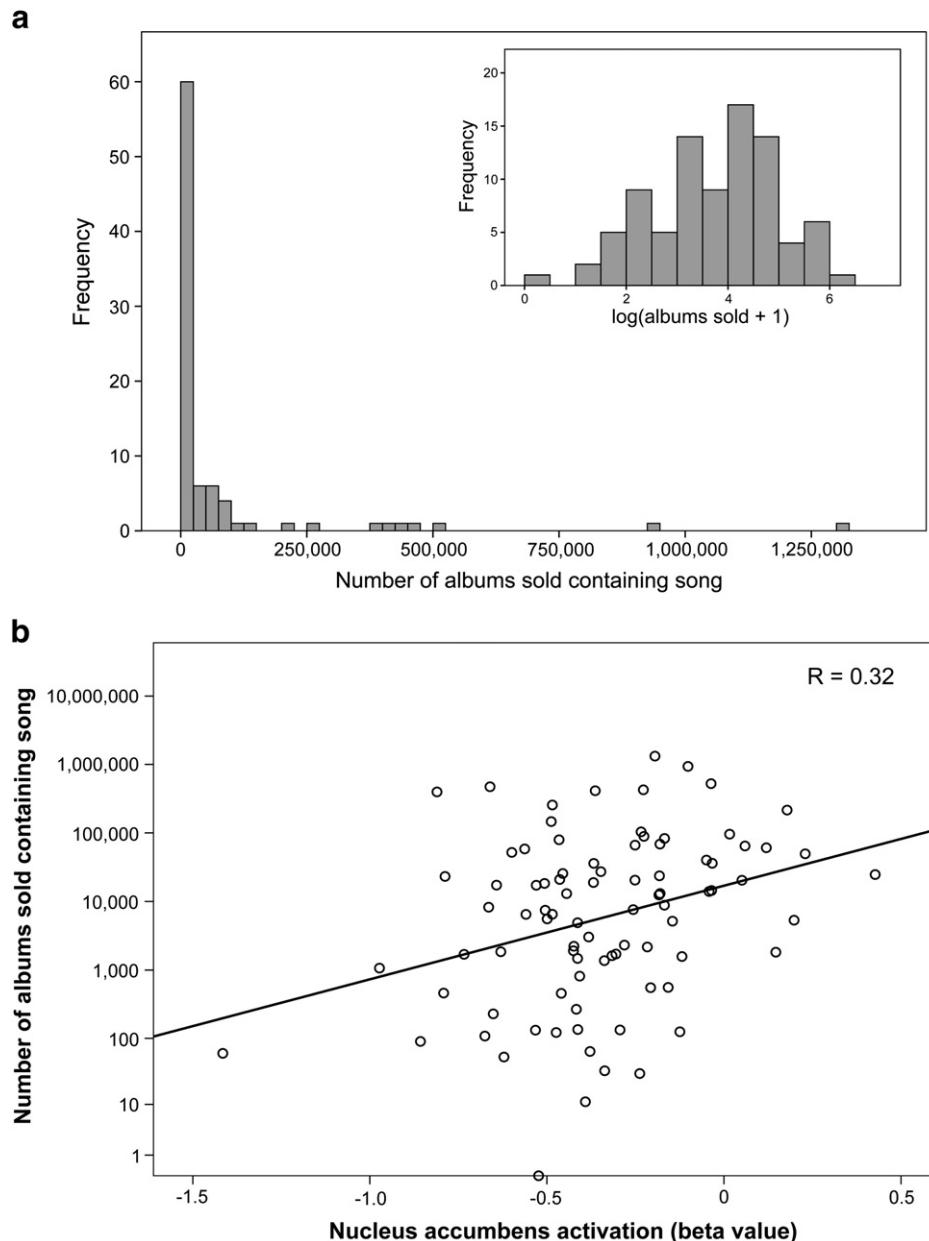


Fig. 2. Distribution of number of albums sold and correlation with nucleus accumbens activation. a) Song sales were determined by sales data reported by Nielsen SoundScan from the album release through May 2010. The distribution was positively skewed, indicating that most songs did not sell many units and with a long tail to the right. Log-transformation of the sales data normalized this distribution for subsequent correlations (*inset*). b) The log of the number of units sold of a song was significantly correlated with the average nucleus accumbens [MNI coordinates: 9, 6, -9; $p < 0.005$, 59 voxels] activation during the listening of the song ($R = 0.32$, $p = 0.004$). Exclusion of the far left outlier resulted in a decreased correlation but which was still significant ($R = 0.27$, $p = 0.013$).

purchase decisions, they are also modestly predictive of population effects. While the nascent field of neuromarketing has made claims to this effect, truly prospective data has been lacking (Ariely & Berns, 2010). Surprisingly, our data suggest some validity to these claims. Why might this be? If the specific cohort of participants in an imaging study is representative of a particular population, then it follows that the results should generalize. When it comes to music, however, it may be difficult to know on which dimensions of a population to match (e.g. age, gender, SES, region). The Recording Industry Association of America (RIAA) estimates that the age range of our cohort

accounts for approximately 20% of music sales (www.riaa.com/keystatistics.php). To test whether the musical tastes of our cohort were representative of the population, we compared our cohort's pre-scan genre rankings to the 2009 Nielsen sales by category and found a significant correlation (Kendall's $\tau = -0.733$, $p = 0.05$; assuming that our hip-hop category is equivalent to Nielsen's R&B category), showing that our cohort had similar tastes as the national population.

Our results also raise the question of why the brain activation was predictive of future sales but the self-reported likability ratings were not. One possibility is that the questions were not

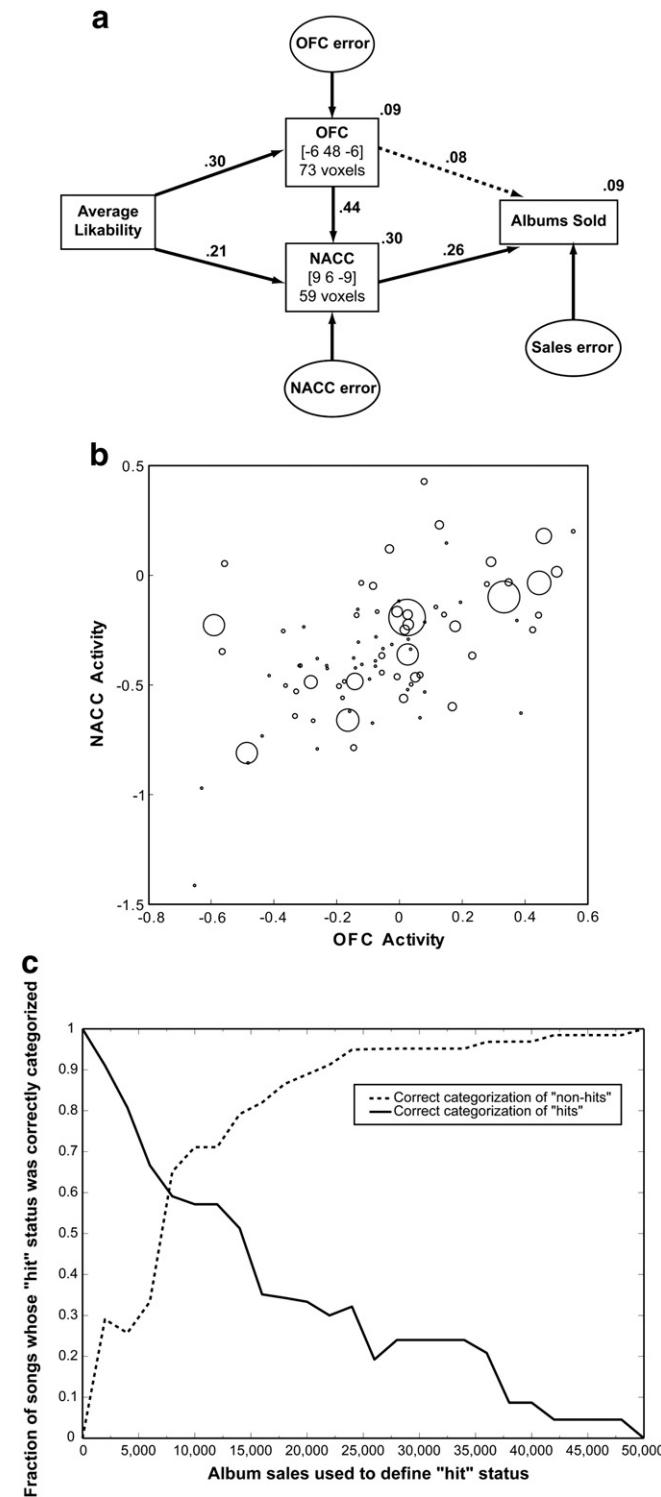


Fig. 3. Structural equation model linking likability ratings to albums sold through OFC and NACC, and logistic regression to categorize "hits." a) Although the average likability of each song was not directly correlated with the number of albums sold, the relationship was moderated by activity in the OFC and NACC. All path coefficients and main effects were significant ($p < 0.05$) except that linking OFC to albums sold. b) The sales of each are represented by the area of each circle and plotted as a function of both OFC and NACC activity. Although the "hits" (big circles) are scattered throughout the activation space, there are a higher proportion of "non-hits" (small circles) toward the lower left. c) To test the ability to correctly categorize "hits" and "non-hits," logistic regressions were performed. The threshold of hit vs. non-hit was varied from 1000 to 50,000 units. Thresholds in the range 15,000 to 35,000 correctly categorize a sizable fraction of hits while also correctly rejecting most non-hits.

adequately specified to differentiate future success (Cronbach & Meehl, 1955). A more detailed debriefing of the reactions to each song might have yielded better predictive data. For example, a choice-based conjoint model might have been superior to simple rankings (Green & Srinivasan, 1990; Griffin & Hauser, 1993). Another possibility is that above a certain quality threshold, songs are too similar to prospectively differentiate them, but slight differences in quality become magnified in superstar markets (Rosen, 1981). Although our results do not invalidate any of these approaches, they do suggest that brain imaging can augment them. The SEM showed that even though likability was not directly correlated with future sales, the OFC and NACC moderated this relationship as essentially hidden variables within the brain. Asking an individual how much they like something requires several cognitive operations, including the initial processing of the stimulus, referencing similar items with which the individual has experience, projection of future utility, all of which may be subject to framing effects of the experiment. In contrast, brain responses in reward-related regions are likely to reflect sub-conscious processes and may yield measurements that are less subject to cognitive strategies. This would be especially true during the consumption of music, which occurred during the listening phase of our experiment. Thus, while the act of rating something requires metacognition, the brain response during the consumption of the good does not, and the latter may prove superior to rating approaches.

Although our data raise the possibility of predicting future popularity in the form of commercial sales, the actual performance of such a model depended on the definition of a "hit." The scarcity of true hits (e.g. 500,000 units) in our sample, underscores the difficulty in evaluating a hit-predictor and confirms the previously noted shift toward superstars (Krueger, 2005). The logistic model performed well in identifying non-hits, which may itself be valuable information, but given the widely varying marketing approaches that are invested in bands (Vogel, 2007), it is surprising that we found any predictive power at all. The fact that we used a wide variety of songs in different genres certainly made the prediction of hits more difficult. A more focused presentation of songs, perhaps within a single genre and pre-screened for minimal quality, would increase the likelihood of hit-prediction. A more targeted group of study participants that is representative of a particular music consuming demographic might also increase predictive power. However, predicting hit-songs may always be a particularly difficult task. A recent study of internet searches found good predictability for revenue of movies and video games but less so for music (Goel, Hofman, Lahaie, Pennock, & Watts, 2010), which makes our results even more surprising and refutes the idea that hits are random (Bielby & Bielby, 1994).

Our results may also have implications for branding. In commercial music, the band is the brand. We calculated commercial success based on the number of units sold, but this number included all sources of a particular song. As a result, the sales numbers were dominated by album sales, which of course contain many songs and may have been heavily influenced by the band/artist reputation (i.e. the band brand). It is hard to know what marketing efforts might have been done to promote a

band. However, it has been estimated that only 10% of new releases end up making a profit for a record label (Vogel, 2007). Consequently, marketing and branding efforts tend to be minimal until a band shows signs of popularity. With more and more artists having access to quality production equipment and being able to release songs directly to the public, neuroimaging tools may have real utility to help labels decide how to invest limited marketing and branding resources.

Acknowledgments

This work was supported by grants from the National Institute for Drug Abuse, the National Science Foundation, the Air Force Office of Scientific Research, and the Office of Naval Research.

References

- Ariely, D., & Berns, G. S. (2010). Neuromarketing: The hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, 11, 284–292.
- Berns, G. S., Capra, C. M., Moore, S., & Noussair, C. (2010). Neural mechanisms of the influence of popularity on adolescent ratings of music. *NeuroImage*, 49, 2687–2696.
- Bielby, W. T., & Bielby, D. D. (1994). "All hits are flukes": Institutionalized decision making and the rhetoric of network prime-time program development. *American Journal of Sociology*, 99(5), 1287–1313.
- Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of Neuroscience*, 29(39), 12315–12320.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., et al. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. *The Journal of Neuroscience*, 26, 6885–6892.
- Goel, S., Hofman, J. M., Lahaie, S., Pennock, D. M., & Watts, D. J. (2010). Predicting consumer behavior with Web search. *Proceedings of the National Academy of Sciences of the United States of America*, 107(41), 17486–17490.
- Green, P. E., & Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, 54(4), 3–19.
- Griffin, A., & Hauser, J. R. (1993). The voice of the customer. *Marketing Science*, 12(1), 1–27.
- Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience*, 28(22), 5623–5630.
- Kenning, P. H., & Plassmann, H. (2008). How neuroscience can inform consumer research. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(6), 532–538.
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, 53, 147–156.
- Krueger, A. B. (2005). The economics of real superstars: The market for rock concerts in the material world. *Journal of Labor Economics*, 23(1), 1–30.
- Lee, N., Broderick, A. J., & Chamberlain, L. (2007). What is neuromarketing? A discussion and agenda for future research. *International Journal of Psychophysiology*, 63, 199–204.
- Lindstrom, M. (2008). *Buyology*. Truth and lies about why we buy. New York: Doubleday.
- Logan, B. R., & Rowe, D. B. (2004). An evaluation of thresholding techniques in fMRI analysis. *NeuroImage*, 22, 95–108.
- McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44, 379–387.
- Montague, P. R., & Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36, 265–284.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4(1), 95–102.
- Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience*, 27(37), 9984–9988.
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences of the United States of America*, 105(3), 1050–1054.
- Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10, 284–294.
- Rosen, S. (1981). The economics of superstars. *American Economic Review*, 71 (5), 845–858.
- Vogel, H. L. (2007). *Entertainment industry economics. A guide for financial analysis* (7th ed.). Cambridge: Cambridge University Press.
- Yoon, C., Gutchess, A. H., Feinberg, F., & Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *Journal of Consumer Research*, 33, 31–40.
- Zhang, H., Nichols, T. E., & Johnson, T. D. (2009). Cluster mass inference via random field theory. *NeuroImage*, 44, 51–61.

Neural Responses to Unattended Products Predict Later Consumer Choices

Anita Tusche,^{1,2} Stefan Bode,^{1,2,3} and John-Dylan Haynes^{1,2}

¹Bernstein Center for Computational Neuroscience Berlin, Charité–Universitätsmedizin Berlin, 10115 Berlin, Germany, ²Max-Planck-Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany, and ³Department of Neurology, Otto-von-Guericke University, 39120 Magdeburg, Germany

Imagine you are standing at a street with heavy traffic watching someone on the other side of the road. Do you think your brain is implicitly registering your willingness to buy any of the cars passing by outside your focus of attention? To address this question, we measured brain responses to consumer products (cars) in two experimental groups using functional magnetic resonance imaging. Participants in the first group (high attention) were instructed to closely attend to the products and to rate their attractiveness. Participants in the second group (low attention) were distracted from products and their attention was directed elsewhere. After scanning, participants were asked to state their willingness to buy each product. During the acquisition of neural data, participants were not aware that consumer choices regarding these cars would subsequently be required. Multivariate decoding was then applied to assess the choice-related predictive information encoded in the brain during product exposure in both conditions. Distributed activation patterns in the insula and the medial prefrontal cortex were found to reliably encode subsequent choices in both the high and the low attention group. Importantly, consumer choices could be predicted equally well in the low attention as in the high attention group. This suggests that neural evaluation of products and associated choice-related processing does not necessarily depend on attentional processing of available items. Overall, the present findings emphasize the potential of implicit, automatic processes in guiding even important and complex decisions.

Introduction

Brain responses obtained during active evaluation of products and explicit deliberation about purchases have been found to predict consumer choices (Knutson et al., 2007). Interestingly, some evidence suggests that automatic brain processes might guide human judgments and choices, even in the absence of explicit deliberation and attention to the choice task. Brain responses were shown to engage automatically in assessing facial attractiveness and preferences even when such judgments were not part of the designated task (O'Doherty et al., 2003; Kim et al., 2007). Likewise, brain activation was reported to reflect preferences when participants evaluate stimuli with respect to other, non-preference-related aspects (Lebreton et al., 2009). However, the precise role of stimulus-related attention in mediating such automatic valuation processes remains unclear. On the one hand, sensory responses to unattended stimuli have been shown to be strongly reduced (Rees et al., 1997; Martínez et al., 1999; Kastner and Ungerleider, 2000). On the other hand, spatially unattended stimuli have been reported to undergo substantial category-

selective processing (Peelen et al., 2009). To date, no study has directly compared neural responses to attended versus unattended products, and the impact of spatial attention on the prediction of economic decisions. Here, we investigated whether brain responses predict consumer choices even when products are entirely task-irrelevant and presented outside the focus of attention.

To examine the role of attention in the prediction of product choices from brain activity, we performed an experiment with two different groups of male participants. In each trial, participants were presented with an image of a car while their brain responses were measured using functional magnetic resonance imaging (fMRI). Participants in group 1 (high attention) were instructed to actively evaluate and rate the attractiveness of each particular car after its presentation (Fig. 1A). Hence, functional brain responses were acquired while products were task-relevant and in the focus of attention. In contrast, participants in group 2 (low attention), were engaged in a demanding visual fixation task while task-irrelevant images of cars were passively presented outside the focus of attention in the background of the screen (Fig. 1B). After the scanning session, participants from both groups were instructed to realistically picture themselves in a consumer setting where they had to decide on a new car. For each of the previously presented products, participants were then asked to state whether they would like to purchase this car or not (Fig. 1C). During scanning, participants from both groups were unaware that they would later be asked about their potential purchases. Multivariate pattern classification (Haxby et al., 2001; Kriegeskorte et al., 2006; Norman et al., 2006; Haynes et al., 2007) was then applied

Received Jan. 6, 2010; revised March 16, 2010; accepted April 8, 2010.

This work was funded by the Max Planck Society, the German Research Foundation, and the Bernstein Computational Neuroscience Program of the German Federal Ministry of Education and Research. We thank Christian Kalberlah, Chun Siong Soon, and Carsten Bogler for their contributions to this project and Thorsten Kahnt for comments on an earlier version of the manuscript.

Correspondence should be addressed to John-Dylan Haynes or Anita Tusche, Charité–Universitätsmedizin Berlin, Bernstein Center for Computational Neuroscience, Haus 6, Philippstrasse 13, 10115 Berlin, Germany. E-mail: haynes@bccn-berlin.de or anita.tusche@bccn-berlin.de.

DOI:10.1523/JNEUROSCI.0064-10.2010

Copyright © 2010 the authors 0270-6474/10/308024-08\$15.00/0

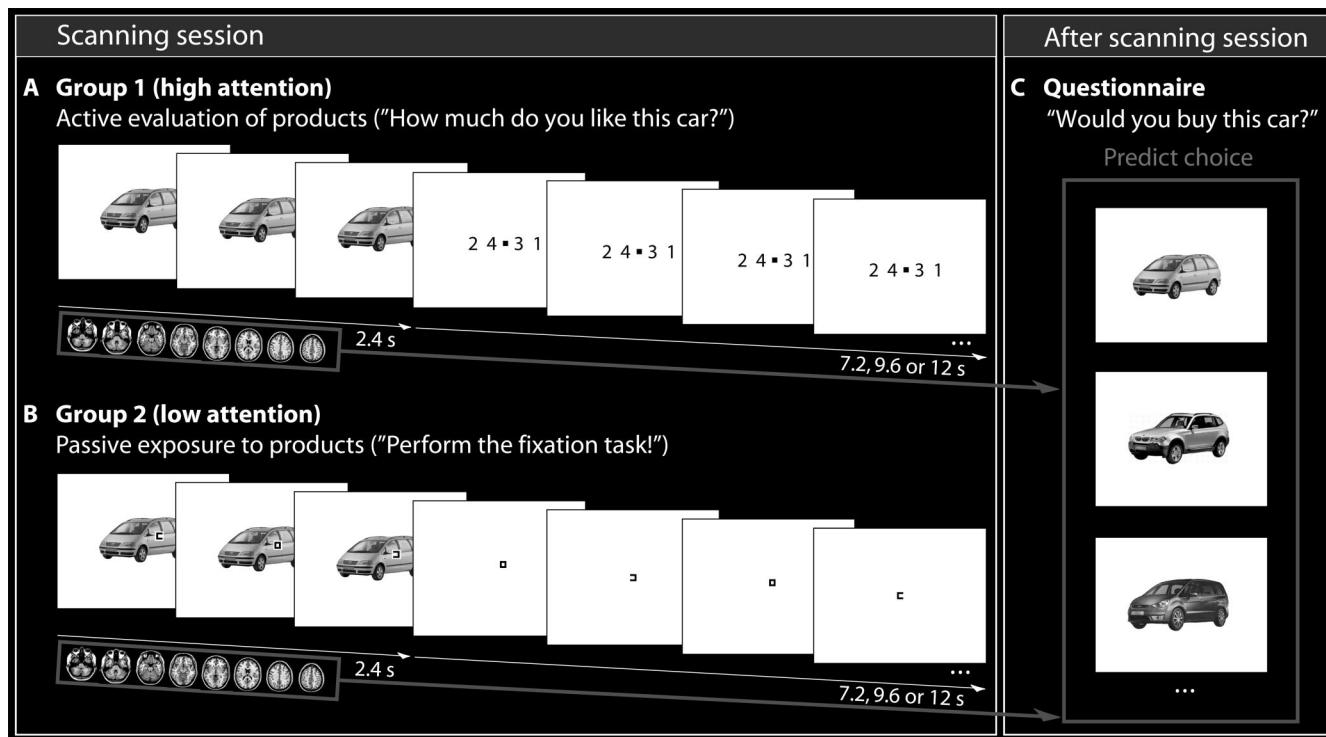


Figure 1. Experimental paradigms. **A**, In each trial, participants in group 1 (high attention) were presented with images of single cars for 2.4 s followed by a randomized response-button mapping screen (displayed for a variable duration of 7.2–12 s). Participants were instructed to closely attend to products and to actively evaluate the attractiveness of the particular car on a four-point scale by pressing the corresponding response button. **B**, Participants in group 2 (low attention) were asked to perform a demanding fixation task. They responded to the opening of a centrally presented fixation square (every 800 ms) with a corresponding left- or right-hand button press. The timing parameters applied in group 1 were retained. Every 7.6–12 s, a car was passively presented on the background of the screen for 2.4 s while the fixation task continued. **C**, After the scanning session, participants from both groups were instructed to realistically picture themselves in a consumer setting where they had to decide upon a new car. Participants were then asked to state their willingness to buy each product. During scanning, participants of both groups were not informed that such a consumer choice would subsequently be required.

to the brain responses obtained during product exposure to predict their stated product choices.

The present study examined the impact of attention on neural predictors of economic decision making. More precisely, we investigated the predictability of consumer choices from brain responses for unattended products. A successful prediction would indicate that product valuation and associated choice-related processing can take place automatically—even with attention diverted from products.

Materials and Methods

Participants and stimuli. Group 1 consisted of 17 participants (aged between 24 and 32 years) and group 2 of 15 participants (aged between 22 and 27 years). Participants were healthy male volunteers, had normal or corrected-to-normal vision, were right-handed and stated that they were interested in cars before the experiment. The sample was limited to males interested in cars to ensure that participants were familiar with and possess stable mental representations about a wide range of items from the product group. Both paradigms were approved by the local ethics committee. All participants were paid €12 to take part and all gave written informed consent.

In both fMRI paradigms, monochrome images of 10 real life cars were used as stimuli. The selection of cars was based on a behavioral pretest with an independent sample of participants in a way that maximized variability in ratings between participants. All pictures were obtained from the Internet and were normalized with regard to size and contrast. During both fMRI paradigms, images were centrally presented against a white background using MATLAB 7.0 (The MathWorks) and the Cogent toolbox (<http://www.vislab.ucl.ac.uk/Cogent>). Data from four participants (in group 1 and 2) as well as data for one car (group 1) and two cars (group 2) was excluded because of missing variance in consumer choices.

Experimental paradigms. During each trial of both event-related fMRI paradigms, participants were presented with a single image of a car for 2.4 s. Participants in group 1 (high attention) were instructed to judge the subjective attractiveness of the specific car on a four-point-scale via a button-press (1 = “dislike it a lot,” 2 = “dislike it a little,” 3 = “like it a little,” 4 = “like it a lot”). The mapping of buttons to attractiveness values was randomized on a trial-by-trial basis to avoid motor preparation during product exposure. Thus, each presentation of a car was followed by a randomized response-mapping screen, which was presented for 7.2–12 s. Responses were given using the index and middle fingers of both hands operating separate button boxes. Participants in group 2 (low attention) were instructed to attend to a black square presented centrally on a white screen. Every 800 ms, the square opened to either the left or the right side. Participants had to respond to each opening with a matching left or right button press using the index fingers of both hands. After a pseudo-randomized duration of 7.2–12 s, a single image of a car was presented for 2.4 s in the background of the screen while the fixation task continued. Thus, the visual presentation of cars in group 2 mirrored that of group 1 but differed in the direction of attention that was diverted from products. Participants in group 2 were explicitly instructed to maximize performance on the fixation task throughout the entire experiment.

For both paradigms, scanning was performed in a single measurement session during which seven independent runs were acquired. The runs were separated by breaks of ~1 min during which no scanning data were obtained. Within a run, each of the 10 cars (see stimuli section above) was presented three times, resulting in a total number of 30 trials per run. For each run, the presentation order of cars was pseudo-randomized such that the same product was never shown in two consecutive trials. Pseudo-randomized durations between presentations of cars varied between 7.2 to 12 s, meaning that each car was combined with different interstimulus

intervals equally often. This procedure ensured that the onset time of the next stimulus was unpredictable and event-related brain responses were clearly separable between trials.

Subsequent to the scanning sessions in both groups, participants were given a questionnaire with the question “Would you buy this car?” and the response options “No/Not sure/Yes.” Importantly, during the acquisition of brain responses participants were unaware that it would be necessary to make such a choice later on. In addition to consumer choices, participants in group 2 had to explicitly judge the attractiveness of each particular car on a four-point scale. Moreover, for both groups familiarity ratings were obtained for all of the products presented. A between-subject design was chosen to ensure that participants in the “low attention” condition were completely unaware of the task-relevance of the products presented.

Image acquisition. For both groups, functional imaging was performed on a 3-Tesla Siemens TRIO scanner with a standard head coil. T2*-weighted functional images were obtained using an echoplanar imaging (EPI) sequence [repetition time (TR) = 2.4 s, echo time = 30 ms]. For each run, 152 EPI volumes were collected (36 ascending axial slices per volume, slice thickness 2 mm, in-plane resolution 3 mm × 3 mm, 1 mm interslice gap, matrix size 64 × 64). The whole session consisted of seven runs. Due to technical problems, only five runs were acquired for two participants in group 2.

Data analysis. Data from both groups were analyzed in a similar manner. In a first step, the acquired volumes were slice timed and realigned. Preprocessed data were then analyzed using a general linear model (GLM) (Friston et al., 1994) as implemented in SPM2 (<http://www.fil.ion.ucl.ac.uk/spm>). For every run, parameter estimates for “purchase” and “no purchase” choices were estimated, based on stated consumer choices obtained after the scanning session. Additionally, biometric pulse data acquired during the scanning were included as coregressors of no interest to account for pulse artifacts. Four regressors were created by using the integral of the pulse curve of four successive time bins (TR/4 = 0.6 s) during the product presentation phase.

In the second step, multivariate pattern classification using a support vector machine (SVM) was applied to the parameter estimates of the consumer choices (Cox and Savoy, 2003; Mitchell et al., 2004; Kamitani and Tong, 2005; Haynes and Rees, 2006). To realize the classification a standard radial basis function kernel as implemented in LIBSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm>) was used. In contrast to a classifier restricted to linear effects, this approach allowed interactions between features and nonlinear functions to drive the prediction of subsequent choices. A comparative figure illustrating prediction accuracies obtained with a linear classification compared with our nonlinear approach is provided in supplemental Figure 1 (available at www.jneurosci.org as supplemental material).

Multivariate pattern classification techniques take advantage of information contained in multiple voxels distributed across space. They allow investigating whether spatial patterns of brain activation contain stable information about different experimental conditions (e.g., purchase vs no purchase). To achieve best predictive accuracy, the classifier weights the contributions of the different voxels optimally. In our case, some voxels within a searchlight cluster were weighted positively and others negatively by the classification algorithm. Moreover, the nonlinear classifier also takes their interactions into account.

To ensure an unbiased analysis of the neural activation patterns throughout the whole brain, a “searchlight” approach was used (Kriegeskorte et al., 2006; Haynes et al., 2007). Given that this approach does not depend on a priori assumptions about informative brain regions or prior voxel selection, the problem of circular analysis (or “double dipping”) can be avoided (Kriegeskorte et al., 2009). For each participant, a sphere with a radius of 4 voxels was created around every voxel v_i of the measured volume. For each sphere, we investigated whether the local pattern of activation during product exposure predicted the willingness to buy that was stated after scanning (purchase vs no purchase). It should be noted that this analysis is predictive because it uses brain activity during exposure to predict purchase ratings obtained after scanning.

For every run, parameter estimates from the GLM were extracted for each of the N voxels in the sphere around voxel v_i and transformed in an

N -dimensional pattern vector. For each run, two pattern vectors were separately created for the purchase and no purchase conditions (see supplemental Fig. 2A, available at www.jneurosci.org as supplemental material). Initially, the pattern vectors of six of the seven runs were used for the training (“training dataset”) of the nonlinear support vector machine classifier with a fixed regularization parameter $C = 1$. This provided the basis of the subsequent classification of the pattern vectors of the remaining seventh run (“test dataset”) as belonging to either the purchase or no purchase condition (see supplemental Fig. 2B, available at www.jneurosci.org as supplemental material). The procedure was independently repeated seven times with a different run used as the test dataset to achieve a sevenfold cross-validation. (A fivefold cross-validation was realized for two participants in group 2 because complete data were acquired only for 5 of the 7 runs). The amount of purchase-related information of the spatial activation pattern of each spherical cluster was represented by the average decoding accuracy across all cross-validation steps and was assigned to the central voxel v_i of the cluster. Given that the number of considered dimensions (equivalent to number of voxels in a searchlight) exceeded the number of acquired data points, we have to consider that noise was fitted to the training of the classifier. The use of independent data for training and testing, however, controlled for the impact of potential overfitting, and consequently for the overestimation in the prediction accuracy (Vul et al., 2009). Moreover, to control for “peaking” problems, decoding accuracies were calculated by estimating the mean across seven cross-validation steps.

The described classification was successively performed for all clusters created around every measured voxel, resulting in a three-dimensional map of average classification accuracies for each participant. These accuracy maps were then spatially normalized to a standard brain [Montreal Neurological Institute (MNI) EPI template as implemented in SPM2] and resampled to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$. Finally, a standard second-level statistical analysis as implemented in SPM2 was performed to identify brain regions that allowed classification of consumer choices across participants in each group. For both groups, analyses were based on the normalized three-dimensional accuracy maps of each subject. Each single point of an individual accuracy map represented the average decoding accuracy of a searchlight surrounding this position across all cross-validation steps. Thus, each value represented the amount of choice-related information of the spatial activation pattern of a surrounding spherical cluster. To assess the statistical significance of these accuracy values across subjects, individual accuracy maps of one group were submitted to a voxelwise one sample t test and contrasted against chance level. Since the classification was based on two alternatives (purchase vs no purchase), chance level was 50%. Familywise error (FWE) correction for multiple comparisons was implemented to control for false positives. Only regions passing this stringent statistical threshold ($p < 0.05$, FWE-corrected, whole brain) and showing significant decoding accuracies above chance were considered relevant for information encoding (Haynes et al., 2007; Soon et al., 2008).

Supplemental data analysis. After each product presentation, participants in group 1 (high attention) had to rate the attractiveness of the particular car on a four-point scale via a button-press. To ensure that brain regions predictive for subsequent consumer choices do not mainly reflect attractiveness, we conducted an additional multivariate decoding analysis. Except for the GLM parameter estimates representing attractiveness ratings instead of consumer choices, the multivariate searchlight decoding was identical to the one described for the main analysis. Another separate decoding analysis was performed on the button-presses to confirm that regions predicting product choices did not simply encode subsequent motor responses. Here, parameter estimates of the GLM were created based on motor responses indicating subjective attractiveness judgments. Apart from that, the analysis was similar to the one described above to predict consumer choices. In both additional decoding analyses chance level was 25%, because the classification was always based on four alternatives of either attractiveness or button-presses. We considered only regions showing significant decoding accuracies above chance as relevant for encoding of attractiveness and motor responses.

Furthermore, classic univariate analyses were conducted for data from both groups. This enabled us to investigate whether there are single vox-

els whose mean activation is significantly more strongly activated in one experimental condition of consumer choices compared with another (purchase vs no purchase). The functional imaging data were preprocessed using slice time correction, motion correction, were spatially normalized to a standard stereotaxic space (MNI EPI template), resampled to an isotropic spatial resolution of $3 \times 3 \times 3 \text{ mm}^3$ and smoothed with a Gaussian kernel of 6 mm full-width half-maximum. Except for these differences in the preprocessing, parameter estimates were created as described in the first step of the multivariate decoding approach. Parameter estimates of purchase and no purchase choices were then contrasted against each other on a single-subject level. Subsequently, a second-level statistical analysis was performed to identify regions that were significantly more strongly activated in one of the two conditions across participants.

To test whether engagement of attention in one region of the visual field (corresponding to the square of the fixation task in group 2) can strongly decrease the processing of task-irrelevant background stimuli, we compared the blood oxygenation level-dependent (BOLD) signal change during product exposure of group 2 (low attention) and group 1 (high attention). In the first step, GLM parameter estimates for visual responses during product exposure were created for each participant and contrasted against baseline. Based on these contrasts, the individual BOLD signal change was estimated for each participant. In the second step, these estimates were used to calculate the average BOLD signal change across participants by implementing a second-level analysis for Groups 1 and 2 individually. Subsequently to contrasting the resulting parameter estimates against baseline, we identified the peak activation area of both hemispheres across these two contrast images (located in the left [MNI $-30, -81, -21$] and right visual cortex [MNI $27, -84, -18$]). Assuming that the visual information was encoded in both hemispheres, we pooled the signal change from both peak activation areas. This was performed for group 1 and group 2 individually. Finally, a *t* test for independent samples was applied to compare the average signal change in group 1 and group 2.

Results

Behavioral results

Consumer choices

In group 1 (high attention), we obtained functional brain responses while participants closely attended to and actively evaluated products. In group 2 (low attention), participants performed a distraction task at fixation while task-irrelevant products were presented outside the focus of attention on the back of the screen. Subsequent to scanning, participants in both groups had to state their willingness to purchase each particular car. The number of trials assigned to either the purchase or no purchase condition was found to be well balanced within both groups and comparable across them. In group 1, the mean distribution of the declared consumer choices for all products was 54% no purchase, 5% “maybe,” and 41% purchase. In group 2, 46% of the products were chosen to be no purchase trials, 4% maybe, and 50% purchase trials across participants. For both groups, profiles of product selection were found to vary across participants (see supplemental Table 1, available at www.jneurosci.org as supplemental material). Additionally, the results of a *t* test for independent samples confirmed that product-specific means (group 1: mean $1.87 \pm \text{SD } 0.83$; group 2: mean $1.86 \pm \text{SD } 0.93$) of consumer choices were comparable across both groups ($t = -0.43, p = 0.67$).

Familiarity

After scanning, participants from both groups were asked whether they had been familiar with the presented products before the experiment. In group 1 (high attention), participants reported the products were familiar in 85% of all cases and as being unknown in 12%. In group 2 (low attention), participants were acquainted with 87% of the products before the experiment

while 10% of products were unknown. For the remaining 3% in both groups the participants stated they were unsure whether they were acquainted with the product before the experiment. For both groups, these findings confirmed that most of the products presented were well known before the experiment.

Task performance and recognition rates

To examine whether attention was effectively removed from products in the low attention condition as implemented in group 2, we analyzed the behavioral performance in the visual fixation task during scanning. The high mean percentage of 89% correct responses indicates that participants attended to the assigned demanding fixation task as instructed. Moreover, we compared the recognition performance in a memory test for both groups (high and low attention) conducted upon completion of the experiment (independent sample of 41 healthy male volunteers: high attention group: $N = 20, 26.5 \pm 2.8$ years; low attention group: $N = 21, 25.8 \pm 1.9$ years, mean \pm SD). Subsequent to the completion of either the high attention or the low attention condition, participants were presented with single images of 20 cars in randomized order. Ten of these cars were stimuli used in the experiment while the remaining 10 images displayed previously unseen cars. For each product, participants had to state whether this car had been presented during the experiment or not. Size and position of the products on the computer screen were kept constant to facilitate recognition performance. Individual hit-rates [HR(high attention) = mean $0.97 \pm \text{SD } 0.07$, HR(low attention) = mean $0.63 \pm \text{SD } 0.26$] as well as false-alarm-rates [FAR(high attention) = mean $0.02 \pm \text{SD } 0.04$, FAR(low attention) = mean $0.22 \pm \text{SD } 0.18$] were found to differ significantly between high and low attention conditions (HR: $t = -5.41, p < 0.001$; FAR: $t = 4.74, p < 0.001$). Participants in the low attention condition showed lower correct recognition of previously seen cars than participants in the high attention condition. Moreover, they were more likely to incorrectly rate a previously unseen car as having been presented during the experiment. These findings strongly suggest that attention was effectively removed from products in the low attention condition of our experiment.

fMRI results

Multivariate decoding of subsequent consumer choices

In group 1 (high attention), spatial activation patterns in the prefrontal cortex (PFC), namely in the left medial frontal gyrus (82% decoding accuracy), the right dorsomedial PFC (75% decoding accuracy) and the bilateral ventromedial PFC (73% decoding accuracy) predicted subsequent consumer choices. Moreover, the left insula (73% decoding accuracy) and the right parahippocampal gyrus (72% decoding accuracy) were found to contain stable information about later product choices (see Table 1 for a complete list of results).

In group 2 (low attention), activation patterns in the left medial PFC (76% decoding accuracy) and the bilateral insula (right: 82% decoding accuracy, left: 72% decoding accuracy) predicted subsequent consumer choices. Neural responses in the left inferior parietal lobe (82% decoding accuracy) and the bilateral superior temporal gyrus (left: 74% decoding accuracy, right: 70% decoding accuracy) also encoded choices between cars (see Table 2 for a complete list of results). It should be noted that decoding accuracies in brain regions predicting subsequent consumer choices under high and low attention conditions were found to be comparable (Fig. 2).

To provide further evidence for the statistical validity of the results obtained by this approach, an additional decoding analy-

Table 1. Brain regions encoding the subsequent consumer choices among actively evaluated products in group 1 (high attention)

Brain region	Side	BA	Accuracy		t value	MNI coordinates		
			M	SE		x	y	z
Frontal lobe								
Medial frontal gyrus	L	9	82	3.3	9.49	-21	27	30
mPFC (dorsal)	R	10	75	2.6	9.53	9	57	21
mPFC (ventral)	L	10	73	2.3	9.96	-9	63	12
Limbic lobe								
Insula	L		73	2.3	9.80	-36	-6	0
Parahippocampal gyrus	R	30	72	1.9	10.98	30	-57	3
Occipital lobe								
Inferior occipital lobe	L	18	73	2.2	10.44	-15	-105	-6
	L	18	70	2.0	9.96	-27	-93	-15
Cerebellum								
	R		77	2.7	9.69	9	-66	-36

Results are reported on a statistical level of $p < 0.05$, FWE-corrected; only peak activations of clusters are listed. L, Left hemisphere; R, right hemisphere; M, mean; BA, Brodmann area.

Table 2. Brain regions encoding the subsequent consumer choices among passively presented, task-irrelevant products in group 2 (low attention)

Brain region	Side	BA	Accuracy		t value	MNI coordinates		
			M	SE		x	y	z
Frontal lobe								
Middle frontal gyrus	L	8	79	2.8	10.45	-24	36	42
	L	8	76	1.8	10.70	-21	36	51
	L	8	72	1.9	11.35	-39	33	42
mPFC	L	10	77	2.2	12.38	-15	57	24
	L	10	76	2.0	12.83	-6	51	18
mOFC	L	11	74	2.0	12.20	-12	42	-18
Inferior frontal gyrus	L	45	73	2.1	10.90	-51	24	15
Limbic lobe								
Insula	R		82	2.6	5.19	39	-3	15
	L		72	1.8	12.46	-42	15	-6
Parietal lobe								
Inferior parietal lobe	L	40	82	2.7	11.81	-57	-36	27
	R	40	73	2.1	10.66	60	-39	30
Temporal lobe								
Superior temporal gyrus	L	38	74	1.8	13.31	-51	12	-18
	R	42	75	1.8	13.40	54	-33	15
	R	42	70	1.5	13.62	63	-36	18

Results are reported on a statistical level of $p < 0.05$, FWE-corrected; only peak activations of clusters are listed. L, Left hemisphere; R, right hemisphere; M, mean; BA, Brodmann area; OFC, orbitofrontal cortex.

sis was conducted for data of both groups. When test datasets were randomly assigned to either the purchase or no purchase condition during the testing phase of the classifier, no statistical significant prediction of consumer choices could be achieved (FDR and FWE-corrected). This was true for data from both groups independently. Moreover, decoding accuracies in brain regions that were informative in the original analysis were at chance level (50%) when the test data were randomly allocated to the conditions (see supplemental Fig. 3, available at www.jneurosci.org as supplemental material). This finding speaks against potential methodological concerns such as possible biases inherent in the testing procedure and insufficient corrections for multiple comparisons. Finally, we investigated whether a combination of two informative classifiers [i.e., medial prefrontal cortex (mPFC) and insula] could improve the overall prediction accuracy. Compared with decoding results of single searchlights in these areas, the weighted classification using decision values of the first-level decoding enhanced the prediction by 7% and by 5% respectively.

Neural activation in the ventral striatum has frequently been implicated in financial decision-making, preference-related processing of products as well as purchases (Erk et al., 2002; Kuhnen and Knutson, 2005; Knutson and Bossaerts, 2007; Knutson et al., 2007; Schaefer and Rotte, 2007a). To investigate whether spatial activation patterns in these regions would be found to contain

stable information about product choices when smaller searchlights (radius of 2 voxels) were used, an additional decoding analysis was performed. At a more liberal statistical threshold of $p < 0.00001$ (uncorrected), this analysis revealed predictive information in the striatum (see supplemental Fig. 4, available at www.jneurosci.org as supplemental material).

Univariate comparisons of subsequent consumer choices

In both groups, classic univariate comparisons did not reveal any activation differences between products that participants were willing to purchase and those they were not. This strongly suggests that multivariate pattern classification is capable of extracting information which conventional analyses fail to detect.

Multivariate decoding of attractiveness judgments

Participants in group 1 were instructed to judge the subjective attractiveness of a particular car after each product presentation. Spatial activation patterns in the right middle frontal gyrus (47% decoding accuracy, [MNI 30, 12, 33]), medial frontal gyrus (43% decoding accuracy, [MNI 15, 33, 45]) and left orbitofrontal cortex (40% decoding accuracy, [MNI -27, 33, 18]) were found to encode attractiveness judgments during product presentation. Activity in the left (51% decoding accuracy, [MNI -15, 24, 30]) and right dorsal anterior cingulate cortex (49% decoding accuracy, [MNI 15, 24, 30]), left (41% decoding accuracy, [MNI -18,

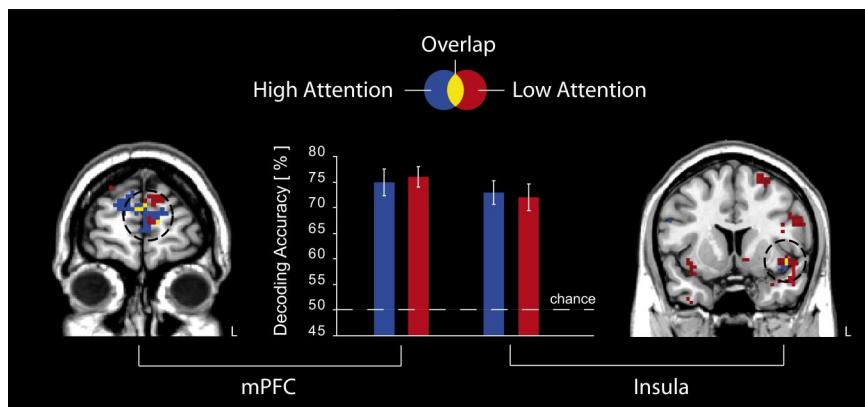


Figure 2. Brain regions encoding subsequent consumer choices in both groups. Multivariate searchlight decoding (radius of 4 voxels) was applied to functional brain responses obtained during product exposure to predict subsequent consumer choices (chance level 50%). Spatial activation patterns in the mPFC and the insula were found to encode these choices when participants did not explicitly deliberate on purchases ($p < 0.05$, FWE-corrected). Importantly, this applied to situations when participants closely attended to and actively evaluated products (“high attention” group 1, blue) as well as when products were passively presented outside the focus of attention (“low attention” group 2, red). The amount of predictive information in the brain responses was found to be comparably high in both groups. The graph displays mean decoding accuracies and SEs across participants for both regions and both groups. For illustrative purposes, the contrasts are shown at $p < 0.0001$ (uncorrected) with a cluster threshold of 10 voxels (L indicates left hemisphere).

−45, 12]) and right posterior cingulate cortex (42% decoding accuracy, [MNI 9, −42, 12]) and left insula (40% decoding accuracy, [MNI −42, 9, −3]) were also found to be predictive of evaluation scores ($p < 0.05$, FWE-corrected, chance level 25% for attractiveness ratings 1–4). For a complete list of results, see supplemental Table 3, available at www.jneurosci.org as supplemental material. Importantly, no direct overlap of regions predicting consumer choices and those predicting attractiveness scores could be observed (see supplemental Fig. 5, available at www.jneurosci.org as supplemental material). This finding strongly supports the notion that the information encoding consumer choices is not merely based on evaluative judgments of attractiveness.

Multivariate decoding of button presses

To ensure that the neural activity during the product viewing phase was not confounded by the preparation of the motor response, we applied a randomized response-button mapping. This mapping scheme was only presented after the removal of the product to isolate the motor response from the processing of product information. As expected, we found no brain region that was informative about the subsequent motor response during the product-viewing phase. This finding strongly suggests that the prediction of subsequent consumer choices was not based on motor preparation during product exposure. During the presentation of the response-mapping scheme (subsequent to the product presentation), the bilateral motor cortex ([MNI 48, −21, 57] and [MNI −42, −21, 54]) predicted the current button-press ($p < 0.05$, FWE-corrected, chance level 25% for buttons 1–4). Importantly, no neural structure predictive of subsequent consumer choices contained information about the associated motor responses (see supplemental Fig. 6, available at www.jneurosci.org as supplemental material). This finding indicates that the prediction of consumer choices was not related to neural processes involved in the execution of motor responses.

Attention modulation in the visual cortex

Engagement of attention in one region of the visual field (corresponding to the square of the fixation task) has been suggested to

decrease the visual processing of task-irrelevant background stimuli (cars). To test this assumption, we compared the BOLD signal change in the visual cortex during product exposure in the high and low attention conditions. Consistent with our hypothesis, we found a significant difference in the BOLD signal change evoked by the presentation of cars in the high attention and low attention group (t test for independent samples, $t = 5.15$, $p < 0.001$). More precisely, the results demonstrated that the responses in the visual cortex evoked by the task-irrelevant images of cars which were presented outside the focus of attention were reduced compared with when the images are actively evaluated and attended to. This finding indicates that the engagement of attention in one region of the visual field (corresponding to the square of the fixation task) decreased the processing of task-irrelevant products in the background.

Discussion

The present study investigated the impact of product-related attention on neural predictors of consumer choices. Activation patterns in the insula and the mPFC were found to predict these choices under high and low attention processing (Fig. 2). Thus, a close match of predictive brain regions was revealed independent of spatial attention to products. This demonstrates that processing of unattended stimuli can proceed beyond object processing (Peelen et al., 2009) to a stage that even allows the prediction of consumer choices. Importantly, the amount of predictive information in these areas was comparably as high when task-irrelevant products were presented outside the focus of attention as when they were actively evaluated and attended to. Congruent with the notion of unconscious environmental triggers for automatic processes in consumer settings (Chartrand, 2005), this indicates that a prediction of economic choices from brain responses does not necessarily depend on attentive processing of products.

Notably, these findings were achieved with participants who were naive about the necessity of a subsequent choice. This is in line with behavioral findings reporting that consumer-related goals can automatically be activated, guiding subsequent consumer behavior and choices even outside of conscious awareness (Bargh, 2002; Chartrand, 2005; Dijksterhuis et al., 2005). Moreover, it is consistent with previous data showing that brain responses reflect preference choices when participants evaluate stimuli with respect to other, non-preference-related aspects (Kim et al., 2007; Lebreton et al., 2009). The current study goes beyond these results by demonstrating that brain responses predict subsequent choices even in the absence of spatial attention to choice options.

Neural activation in the insula and the mPFC has been shown to predict consumer choices when participants closely attend to products and explicitly deliberate about purchases (Knutson et al., 2007). Given that the prediction of consumer choices in the present study was achieved in the absence of explicit deliberation and without a priori assumptions about informative regions, this further supports the role of both areas in economic decision-making. Spatial activation patterns in the insula and the mPFC

were also reported to predict behavioral choices in the following trial of a reward-based decision-making task (Hampton and O'Doherty, 2007). In line with this finding, both structures have been suggested to underlie the neural representation of the expected reward value (Knutson et al., 2005; Preuschoff et al., 2006; Hampton et al., 2007; Hare et al., 2008; Rolls et al., 2008), a concept central to microeconomic and psychological models of decision-making (Knutson and Bossaerts, 2007; Loewenstein et al., 2008; Rangel et al., 2008). Cultural stimuli such as cars and logos of car manufacturers can signal potential social dominance or wealth and have been shown to modulate activity in the reward circuit (Erk et al., 2002; Schaefer and Rotte, 2007a). It can be assumed that products in the present experiment differed in their expected reward value depending on whether they were subsequently chosen to be purchased or not (Sharot et al., 2009). We suggest that distinguishable activation patterns in the insula and the mPFC reflected these different reward values and—even without full attention—contributed to subsequent hypothetical consumer choices.

Activation in the mPFC—particularly the ventral part—has also been reported to reflect product-related preference and attractiveness judgments (Paulus et al., 2003; McClure et al., 2004; Deppe et al., 2005; Plassmann et al., 2008; Lebreton et al., 2009; Luu and Chau, 2009). The dorsal mPFC, on the other hand, has been proposed to be involved in the processing of brand knowledge (Schaefer et al., 2006; Schaefer and Rotte, 2007b), the impact of which on consumer preferences and choices is well known (McClure et al., 2004; Lee et al., 2006). Considering this evidence, it is likely that the predictive information encoded in the mPFC might have been influenced by subjective valuation and the brand and price information.

It can be assumed that global product valuation was a major source for subsequent choices. This is particularly likely, given that no further product-related information was provided during the experiment to avoid drawing participants' attention to the potential choice options. Therefore, it could be speculated that the predictive information merely reflects global attractiveness or desirability of products. However, many cognitive factors as well as their interactions with automatic valuation processes might contribute to a complex choice such as those for a new car. This is particularly likely for participants who stated an interest in cars and are assumed to possess relevant knowledge on the items. Consistent with this notion product choices were not entirely determined by attractiveness judgments. Thus, only partial correlations between product-specific attractiveness ratings and consumer choices were found for both groups in the present experiment (see supplemental Table 2, available at www.jneurosci.org as supplemental material). Additionally, no direct overlap of brain regions predicting attractiveness judgments and those being informative about consumer choices could be identified (see supplemental Table 3 and supplemental Fig. 5, available at www.jneurosci.org as supplemental material). Together, these findings indicate that the prediction of product choices was not mainly due to global evaluations of attractiveness but might reflect automatic choice-related processing itself. This is in line with previous results showing that brain activation reflecting subsequent preference decisions were not merely responding to attractiveness of stimuli (Lebreton et al., 2009). However, more research is needed to specifically examine choice-relevant dimensions with the current paradigms, possibly testing explicit models of multiattribute decision-making (Dijksterhuis et al., 2006; Lassiter et al., 2009). Given that consumer choices for each product were found to vary from participant to participant, it is also unlikely that the

predictive information was based on physical properties of single cars (see supplemental Table 1, available at www.jneurosci.org as supplemental material). Results of an additional decoding analysis also demonstrate that the prediction of consumer choices was not related to neural processes involved in the preparation of motor responses (see supplemental Fig. 6, available at www.jneurosci.org as supplemental material).

Another point that needs to be addressed concerns the question of whether attention was effectively attenuated by engaging in the distraction task in group 2. Evidence of successful distraction comes from differences in memory performance, obtained for both groups after the experiment. Recognition rates of cars were strongly decreased in participants who performed the distraction task compared with participants who actively evaluated products. Moreover, neural responses in the visual cortex evoked by task-irrelevant, unattended images of cars were found to be reduced compared with images that were actively evaluated and fully attended to. This is consistent with previous research demonstrating that the engagement of attention in one region of the visual field strongly decreases the processing of irrelevant background stimuli (Rees et al., 1997). It can therefore be considered that the attention to products in group 2 was strongly diminished compared with group 1. As in most attention studies, there is a possibility of weak but residual attention to unattended target stimuli. However, our key finding is that a strong reduction in attention does not affect the choice-predictive information. Purchase-related information encoded in the brain was not reduced, but was comparably high in low and high attention conditions. Furthermore, we found a close match of predictive brain regions during high and low attentional processing of products. Together, this indicates that choice-related processing does not necessarily require close attention to products and can even occur in cases where sensory signal processing is attenuated due to a removal of attention. Although it remains possible that the choice was not explicitly calculated during product presentation, the high predictive accuracy reveals that choice-related processes have already reached a high level of completion.

The present study implemented hypothetical choices, which are commonly applied in marketing research. Nevertheless, it needs to be explored whether the same activation patterns are predictive for actual purchases—beyond the stated willingness to buy. This would allow checking for potential biases such as the tendency to overstate the willingness to pay or vote in case of hypothetical choices (List and Gallet, 2001; Murphy et al., 2005). Requiring expenses would likely involve losses and more relative comparisons between available products, including changes in the reference point for product valuation (FitzGerald et al., 2009; Sharot et al., 2009). Real purchases might also engage processes of perceived justification, anticipated regret, time pressure or self esteem (Plassmann et al., 2007). Finally, it might be that making actual purchases differs from our experimental setting in terms of the strategies used for information acquisition and information integration across multiple relevant dimensions. However, it should be noted that informative brain regions as identified with the present approach are strikingly consistent with previous findings that implemented actual purchases (Knutson et al., 2007). Further research might also address the generalizability of our findings to other product categories, including different goods (e.g., cars vs coffee) and different types of purchases (e.g., routine vs new), as well as to people who are not interested in the product group.

In summary, we found a close match of brain regions predicting consumer choices for both high and low attentional processing of products. Importantly, the amount of predictive information was

found to remain persistently high when task-irrelevant products were presented outside the focus of attention. Altogether, these findings support the notion that even complex and important economic choices can be prepared automatically, in the absence of explicit deliberation and without attention to products.

References

- Bargh JA (2002) Losing consciousness: automatic influences on consumer judgment, behavior, and motivation. *J Consum Res* 29:280–285.
- Chartrand TL (2005) The role of conscious awareness in consumer behavior. *J Consum Psychol* 15:203–210.
- Cox DD, Savoy RL (2003) Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19:261–270.
- Deppe M, Schwindt W, Kugel H, Plassmann H, Kenning P (2005) Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J Neuroimaging* 15:171–182.
- Dijksterhuis A, Smith PK, van Baaren RB, Wigboldus DHJ (2005) The unconscious consumer: effects of environment on consumer behavior. *J Consum Psychol* 15:193–202.
- Dijksterhuis A, Bos MW, Nordgren LF, van Baaren RB (2006) On making the right choice: the deliberation-without-attention effect. *Science* 311:1005–1007.
- Erk S, Spitzer M, Wunderlich AP, Galley L, Walter H (2002) Cultural objects modulate reward circuitry. *Neuroreport* 13:2499–2503.
- FitzGerald THB, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci* 29:8388–8395.
- Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1994) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210.
- Hampton AN, O'Doherty JP (2007) Decoding the neural substrates of reward-related decision making with functional MRI. *Proc Natl Acad Sci U S A* 104:1377–1382.
- Hampton AN, Adolphs R, Tyszka MJ, O'Doherty JP (2007) Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron* 55:545–555.
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28:5623–5630.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
- Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7:523–534.
- Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE (2007) Reading hidden intentions in the human brain. *Curr Biol* 17:323–328.
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685.
- Kastner S, Ungerleider LG (2000) Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci* 23:315–341.
- Kim H, Adolphs R, O'Doherty JP, Shimojo S (2007) Temporal isolation of neural processes underlying face preference decisions. *Proc Natl Acad Sci U S A* 104:18253–18258.
- Knutson B, Bossaerts P (2007) Neural antecedents of financial decisions. *J Neurosci* 27:8174–8177.
- Knutson B, Taylor J, Kauffman M, Peterson R, Glover G (2005) Distributed neural representation of expected value. *J Neurosci* 25:4806–4812.
- Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G (2007) Neural predictors of purchases. *Neuron* 53:147–156.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868.
- Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540.
- Kuhnen CM, Knutson B (2005) The neural basis of financial risk taking. *Neuron* 47:763–770.
- Lassiter GD, Lindberg MJ, González-Vallejo C, Bellezza FS, Phillips ND (2009) The deliberation-without-attention effect: evidence for an artifactual interpretation. *Psychol Sci* 20:671–675.
- Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64:431–439.
- Lee L, Frederick S, Ariely D (2006) Try it, you'll like it: the influence of expectation, consumption, and revelation on preferences for beer. *Psychol Sci* 17:1054–1058.
- List JA, Gallet CA (2001) What experimental protocol influences disparities between actual and hypothetical stated values. *Environ Res Econ* 20:241–254.
- Loewenstein G, Rick S, Cohen JD (2008) Neuroeconomics. *Annu Rev Psychol* 59:647–672.
- Luu S, Chau T (2009) Decoding subjective preference from single-trial near-infrared spectroscopy signals. *J Neural Eng* 6:016003.
- Martinez A, Anello-Vento L, Sereno MI, Frank LR, Buxton RB, Dubowitz DJ, Wong EC, Hinrichs H, Heinze HJ, Hillyard SA (1999) Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nat Neurosci* 2:364–369.
- McClure SM, Li J, Tomlin D, Cypert KS, Montague LM, Montague PR (2004) Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44:379–387.
- Mitchell TM, Hutchinson R, Niculescu RS, Pereira F, Wang X, Just M, Newman S (2004) Learning to decode cognitive states from brain images. *Machine Learn* 57:145–175.
- Murphy J, Allen PG, Stevens TH, Weatherhead D (2005) A meta-analysis of hypothetical bias in stated preference valuation. *Environ Res Econ* 30:313–325.
- Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430.
- O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ (2003) Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41:147–155.
- Paulus MP, Rogalsky C, Simmons A, Feinstein JS, Stein MB (2003) Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* 19:1439–1448.
- Peelen MV, Li FF, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460:94–97.
- Plassmann H, Ambler T, Brautigam S, Kenning P (2007) What can advertisers learn from neuroscience? *Int J Advertising* 26:151–175.
- Plassmann H, O'Doherty J, Shiv B, Rangel A (2008) Marketing actions can modulate neural representations of experienced pleasantness. *Proc Natl Acad Sci U S A* 105:1050–1054.
- Preuschoff K, Bossaerts P, Quartz SR (2006) Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51:381–390.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556.
- Rees G, Frith CD, Lavie N (1997) Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science* 278:1616–1619.
- Rolls ET, McCabe C, Redoute J (2008) Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb Cortex* 18:652–663.
- Schaefer M, Rotte M (2007a) Favorite brands as cultural objects modulate reward circuit. *Neuroreport* 18:141–145.
- Schaefer M, Rotte M (2007b) Thinking on luxury or pragmatic brand products: brain responses to different categories of culturally based brands. *Brain Res* 1165:98–104.
- Schaefer M, Berens H, Heinze HJ, Rotte M (2006) Neural correlates of culturally familiar brands of car manufacturers. *Neuroimage* 31:861–865.
- Sharot T, De Martino B, Dolan RJ (2009) How choice reveals and shapes expected hedonic outcome. *J Neurosci* 29:3760–3765.
- Soon CS, Brass M, Heinze HJ, Haynes JD (2008) Unconscious determinants of free decisions in the human brain. *Nat Neurosci* 11:543–545.
- Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Percept Psychol Sci* 4:274–290.

Subliminal Instrumental Conditioning Demonstrated in the Human Brain

Mathias Pessiglione,^{1,2,*} Predrag Petrovic,¹ Jean Daunizeau,¹ Stefano Palminteri,² Raymond J. Dolan,¹ and Chris D. Frith¹

¹Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK

²Laboratoire INSERM U610, Centre de Neuroimagerie de Recherche (CENIR), Institut Fédératif de Recherche en Neurosciences, Hôpital Pitié-Salpêtrière, Université Pierre et Marie Curie (Paris 6), 47 Boulevard de l'Hôpital 75013 Paris, France

*Correspondence: mathias.pessiglione@gmail.com

DOI 10.1016/j.neuron.2008.07.005

SUMMARY

How the brain uses success and failure to optimize future decisions is a long-standing question in neuroscience. One computational solution involves updating the values of context-action associations in proportion to a reward prediction error. Previous evidence suggests that such computations are expressed in the striatum and, as they are cognitively impenetrable, represent an unconscious learning mechanism. Here, we formally test this by studying instrumental conditioning in a situation where we masked contextual cues, such that they were not consciously perceived. Behavioral data showed that subjects nonetheless developed a significant propensity to choose cues associated with monetary rewards relative to punishments. Functional neuroimaging revealed that during conditioning cue values and prediction errors, generated from a computational model, both correlated with activity in ventral striatum. We conclude that, even without conscious processing of contextual cues, our brain can learn their reward value and use them to provide a bias on decision making.

INTRODUCTION

Humans frequently invoke an argument that their intuition can result in a better decision than conscious reasoning. Such assertions may rely on subconscious associative learning between subliminal signals present in a given situation and choice outcomes. For instance, clinicians may improve their therapeutic decisions through learned associations between treatment outcomes and subliminal signs presented by their patients. Likewise, poker players can improve their gambles through a learned association between monetary outcomes and subliminal behavioral manifestations of their opponents (the so-called “gamblers’ tell”).

The idea that such instrumental conditioning can occur subconsciously has been around for almost a century (Thorndike, 1911). This assumption originally rested on observations that rewards and punishments shape behavioral responses in species allegedly lacking conscious awareness. However, subliminal

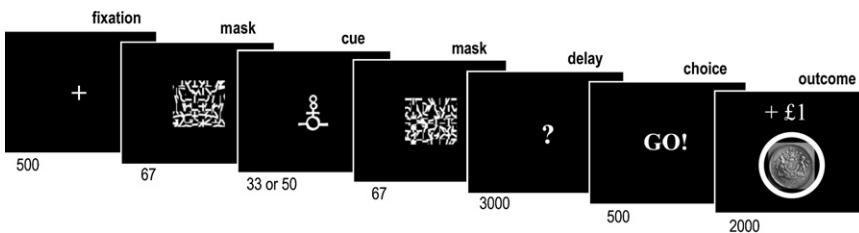
conditioning studies in humans have so far been restricted to Pavlovian paradigms such as fear conditioning (Clark and Squire, 1998; Knight et al., 2003; Morris et al., 1998; Olsson and Phelps, 2004), where subliminal stimuli (like unseen faces) are paired with unpleasant events (like white noise) to increase automatic responses (like skin conductance). To our knowledge, subliminal instrumental conditioning, where decision making would be biased by unperceived cues predicting rewards or punishments, has never been previously demonstrated.

Our subliminal conditioning task was adapted from a published supraliminal task, wherein subjects selected between visual cues so as to learn choices that maximized monetary outcomes (Pessiglione et al., 2006). In our previous study, we modeled subjects’ behavior by optimizing the free parameters of a standard machine learning algorithm (termed Q learning), to get maximal likelihoods for the observed decisions. When we regressed key output variables of the optimized model against simultaneously acquired functional neuroimaging data, we showed that prediction errors were expressed in the striatum. Postexperimental debriefing indicated that some subjects managed to understand the statistical structure of the task, while others appeared to rely on what they referred to as their intuition. These latter reports suggest that subjects can improve their decisions without consciously following the incremental steps of the Q-learning procedure.

The motivating assumption of the current experiment was that processes associated with striatal learning are not consciously accessible but, nonetheless, influence choice decision making. Indeed, if contextual cues reach awareness, other brain systems are likely to play a role, as expressed in conscious reasoning or strategic control, which allows one to develop explicit knowledge of statistical contingencies. However, if the cues remain unseen, learning would solely depend on a subconscious processing that involves the striatum, with an algorithmic structure similar to a Q learning, which does not embody explicit information about statistical contingencies. Under these assumptions, we predicted that, if in our task visual cues were masked, both striatal activity and behavioral choices would still reflect Q-learning outputs.

RESULTS

A prerequisite for the present study was to demonstrate efficient masking of the visual cues. These cues were novel abstract symbols, which were scrambled and mixed to create mask images.

**Figure 1. Subliminal Conditioning Task**

Successive screenshots displayed during a given trial are shown from left to right, with durations in milliseconds. After seeing a masked contextual cue flashed on a computer screen, subjects choose to press or not press a response button and subsequently observe the outcome. In this example, "Go" appears on the screen because the subject has pressed the button, following the cue associated with a rewarding outcome (winning £1).

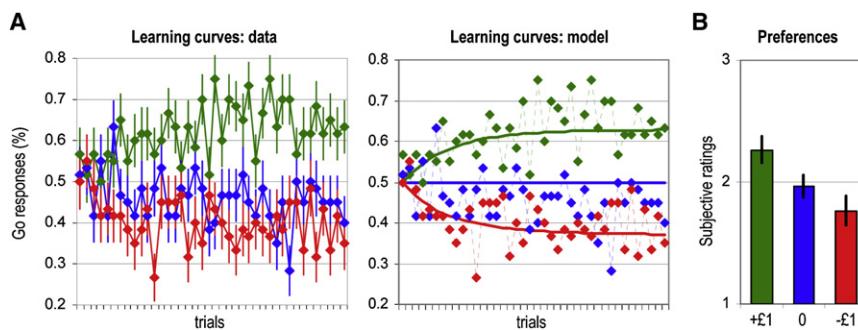
To assess visual awareness, we successively displayed two masked cues on a computer screen and asked subjects whether they perceived a difference or not. We reasoned that if subjects are unable to correctly perceive any difference between the masked cues, then they are also unable to build conscious representations of cue-outcome associations. The procedure has the advantage of not showing the cues unmasked, so that, by the end of the experiment, subjects had no idea what the cues look like.

The perceptual discrimination task was performed outside the scanner at the beginning of the experiment, in order to adapt duration of cue display to each individual, and in the scanner at the end of the experiment, to check for any effect of learning or change in visual conditions. For all subjects, cue duration was set at either 33 or 50 ms and was kept fixed through the entire experiment. In every individual, correct guessing on the final assessment did not differ from chance ($p > 0.05$, chi-square test). At group level, average percentage of correct responses for the 20 subjects was $48\% \pm 3\%$, which again was not different from chance ($p > 0.5$, two-tailed paired t test). Average d' was 0.08 ± 0.20 , showing that, even when correcting for response bias, signal detection was not different from zero ($p > 0.5$, two-tailed paired t test). Thus, subjects remained unable to discriminate between the different masked cues, from the beginning to the end of conditioning sessions.

We employed the same masking procedure in the subliminal conditioning task, in which cues were paired with monetary outcomes (Figure 1). From these outcomes ($-£1$, $£0$, $+£1$), subjects had to learn when to take the risky response (either "Go" or "NoGo," depending on subjects). Subjects were also told that, for the risky response, the outcome would depend on the cue hidden behind the masking image (see instructions in Supple-

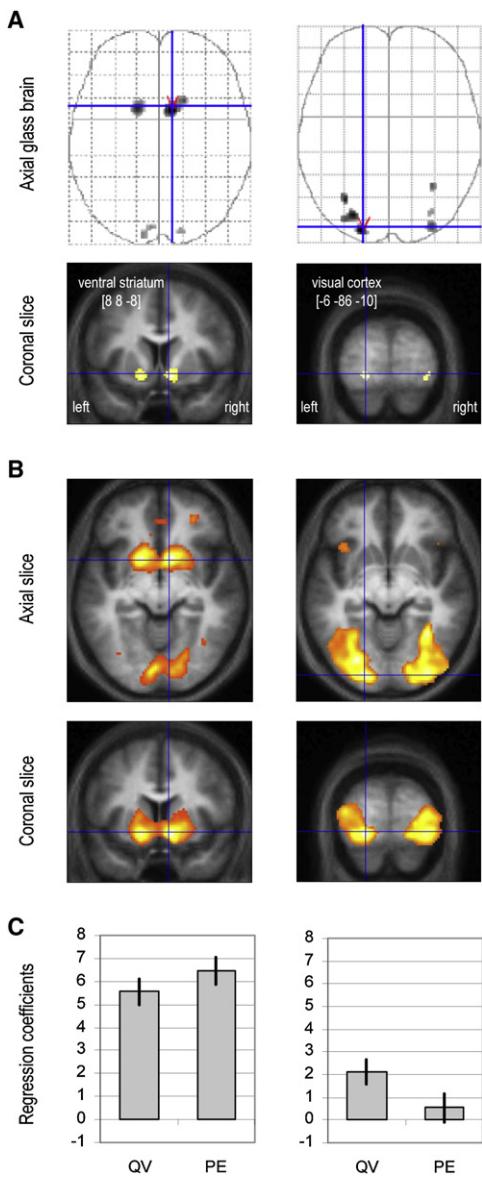
mental Data available online). As they would not see the cues, we encouraged them to follow their intuition, taking the risky response if they had a feeling they were in a winning trial and choosing a safe response if they felt it was a losing trial. Note that if subjects always made the same response, or if they performed at chance, their final payoff would be zero.

As a dependent variable to assess for conditioning effects, we used monetary payoff, which corresponds to the area below the reward and above the punishment learning curves (Figure 2A). Overall subjects won money in this task, on average $£7.5 \pm £1.8$ ($p < 0.001$, one-tailed paired t test), indicating that the risky response was more frequently chosen following reward predictive relative to punishment predictive cues. Both reward and punishment conditions also differed significantly from the neutral condition ($p < 0.05$, one-tailed paired t test). There was no significant difference ($p > 0.5$, two-tailed paired t test) between the reward and punishment condition: on average subjects won $£24.3 \pm £1.9$ and avoid losing $£23.2 \pm £2.1$. Learning curves showed that responses improved symmetrically for rewards and punishments, ending with $63\% \pm 5\%$ of correct responses on average. Surprisingly, this plateau was reached at around the halfway point of the learning session. The effects of subliminal conditioning were subsequently assessed with a preference judgment task, in which cues were uncovered and rated by the subjects from the most to least liked (Figure 2B). Ratings were significantly higher for reward compared to punishment cues ($p < 0.01$, one-tailed paired t test), consistent with subjects having learned the affective values of subliminal cues, such that these values were able to bias their preferences. When uncovering the cues, subjects were also asked to signal any cue that they may have seen during conditioning sessions; none was reported as previously seen.

**Figure 2. Behavioral Data**

(A) Learning curves. Colors indicate cues for which button presses are rewarded (green), neutral (blue), or punished (red). Diamonds represent, across trials, percentages of subjects that pressed the button. Left: continuous lines join the diamonds to illustrate actual choices made by subjects. Right: continuous lines represent the probabilities of button press estimated by an optimized Q-learning model.

(B) Preferences. After the conditioning phase, cues were unmasked and subjects rated them, from the most (3) to the least liked (1). The graph shows the average rating for reward (green), neutral (blue), and punishment (red) cues. Bars are \pm intersubjects standard errors of the mean.

**Figure 3. Neuroimaging Data**

Left: ventral striatal regions isolated by regression of monetary values against BOLD responses to cue display. Right: visual cortical regions isolated by correlation of cue-value regression coefficients with individual payoffs. Slices were taken at global maxima of interest indicated by red pointers on the above axial glass brains. Areas shown in gray/black on glass brains and in orange/yellow on coronal slices showed significant effect. The [x y z] coordinates of the maxima refer to the Montreal Neurological Institute space.

- (A) Statistical parametric maps using conservative threshold ($p < 0.05$ after familywise error correction for multiple comparisons).
- (B) Statistical parametric maps using liberal threshold ($p < 0.001$ uncorrected).
- (C) Regression coefficients of Q values (QV) and prediction errors (PE) against BOLD responses to cue and outcome display, respectively. Bars are \pm inter-subjects standard errors of the mean.

To model instrumental conditioning, we implemented a standard Q-learning algorithm (Pessiglione et al., 2006), with inputs from individual histories of cues, choices, and outcomes. On

every trial, the model estimates the likelihood of the risky response from the value of the displayed cue. If the risky response was actually taken, the model then updates the value of the displayed cue in proportion to the prediction error. The parameters of the model were optimized such that likelihoods of risky responses provided the best fit of subjects' actual responses across conditioning sessions (Figure 2A). The Q values and prediction errors generated by this optimized algorithm were then used as regressors for analysis of brain imaging data (see Figure S1).

We recorded brain activity while subjects performed the subliminal conditioning task, using functional magnetic resonance imaging (fMRI). We first examined brain regions reflecting Q value at the time of cue onset, increasing their response to reward-predicting cues and decreasing their response to punishment-predicting cues, across learning sessions. After correction for multiple comparisons (family-wise error, $p < 0.05$), we noted significant correlated activity in ventral striatum bilaterally (Figures 3A and 3B, left). The same region was also significantly activated at the time of outcome in keeping with prediction errors being expressed at this time point (Figure 3C, left). In a second analysis, we computed regression coefficients for the different conditions at the time of cue and outcome onsets, separately for the first and second half of conditioning sessions. Contrasts with the neutral condition were then averaged over all ventral striatal voxels showing significant activation at the most conservative threshold in the first analysis. This confirmed that from the first to the second half of conditioning sessions, ventral striatal responses increased for reward cues and decreased for punishment cues (Figure 4A, left). At the time of outcome onset, the same ventral striatal region reflected positive prediction errors in the reward condition and negative prediction errors in the punishment condition. In keeping with the Q-learning model, both positive and negative prediction errors decreased from the first to the second half of conditioning sessions. Thus, across subliminal conditioning, the ventral striatal response was consistent with the expression of Q values (for unseen cues) and prediction errors (based on visible outcomes).

We further examined variability in individual performance to explain why some subjects won more money than others. More precisely, we searched for brain regions where coefficients of Q-value regressors correlated with individual payoffs. These regions were confined to extrastriate visual cortex (Figure 3A, right) at the most conservative threshold (familywise error, $p < 0.05$), spreading into the ventral occipitotemporal stream (Figure 3B, right) with a more liberal threshold (uncorrected, $p < 0.001$). Contrast estimates confirmed that extrastriate voxels progressively differentiated the reward and punishment cues from the first to the second half of conditioning sessions (Figure 4A, right). At the time of outcome onset, these extrastriate regions responded positively for both rewards and punishments, showing no evidence for encoding of prediction errors. Thus, during the subliminal conditioning task, the extrastriate visual cortex learned to discriminate between unseen cues according to their reward value but did not express outcome-related prediction errors (Figure 3C, right).

To further assess whether the ventral striatum and visual cortex were able to discriminate between the subliminal cues,

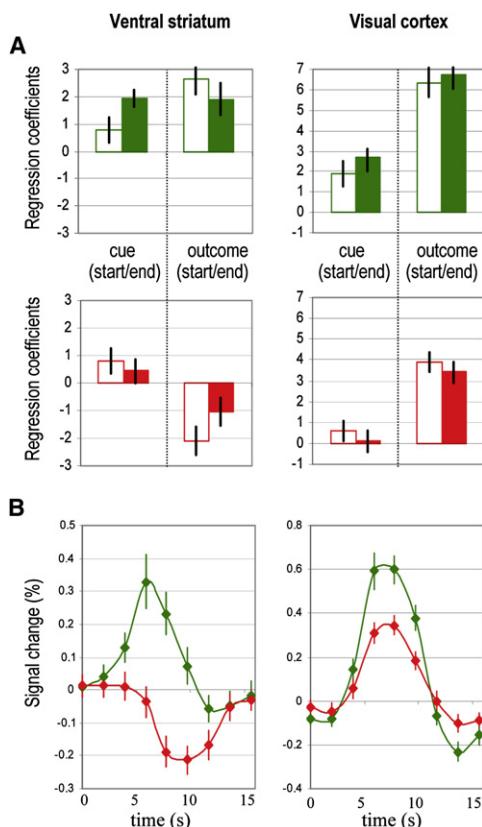


Figure 4. Model-free Analyses of Brain Activations

Ventral striatum (left) and visual cortex (right) correspond to voxels surviving familywise error correction on statistical parametric maps.

(A) Regression coefficients. Histograms represent contrasts of the reward (green) or the punishment (red) condition with the neutral condition, at the times of cue and outcome display. For every contrast the two joint histograms correspond to the first (empty) and second (filled) halves of the conditioning sessions.

(B) Time courses. BOLD responses were averaged across trials over conditioning sessions, for the reward (green) and punishment (red) conditions, relative to the neutral condition. Bars are \pm intersubjects standard errors of the mean.

we extracted time courses of BOLD response. These time courses were averaged over trials, sessions, and subjects, separately for the reward and punishment conditions (Figure 4B). We found that BOLD responses to reward and punishment cues significantly differed after two acquisition volumes (3.9 s) in the ventral striatum (one-tailed paired t test, $p < 0.01$) and after three (5.85 s) in the visual cortex (one-tailed paired t test, $p < 0.01$).

Finally, we ascertained whether neuroimaging and behavioral effects of subliminal conditioning were driven by subjects scoring at the high end in perceptual discrimination performance. We tested for correlations between correct guessing assessed in the final perceptual discrimination test and coefficients of Q-value regressors in both the ventral and extrastriate cortex. None was significant; Pearson's correlation coefficients were respectively -0.25 and -0.18 . We also tested correlation of correct guessing with monetary payoffs from conditioning sessions and differential ratings in the preference judgment task. Again,

none was significant; Pearson's correlation coefficients were respectively 0.26 and 0.29 .

DISCUSSION

We provide evidence that instrumental learning can occur without conscious processing of contextual cues. This finding might relate to previous evidence for implicit or procedural learning, where behavioral responses can be adapted to the statistical structure of stimuli that fails to be reported explicitly (Bayley et al., 2005; Berns et al., 1997; Destrebecqz and Cleeremans, 2001; Seitz and Watanabe, 2003). Interestingly, implicit/procedural learning has been suggested to involve the basal ganglia, in contrast with explicit/declarative memory which would involve the medial temporal lobe (Cohen et al., 1985; Milner et al., 1998; Poldrack and Packard, 2003; Squire, 1992). In implicit learning tasks, such as serial reaction time or probabilistic classification, authors have claimed that subjects can achieve good acquisition without explicit knowledge of the task structure. However, methods for assessing awareness of statistical contingencies have been criticized, principally on the issue that questions were too demanding in terms of memory (Lagnado et al., 2006; Lovibond and Shanks, 2002; Wilkinson and Shanks, 2004). Thus, to formally test whether instrumental conditioning can occur without awareness, we took a more stringent approach: masking the cues, so that they remained unperceived.

We believe our methodology avoids most previous problems related to assessing awareness, by demonstrating that subjects were not able to discriminate between masked cues (without the help of rewards and punishments), rather than retrospectively assessing awareness of contingencies. Moreover, postconditioning recognition tests would not be sufficient in our case, since subjects would not need to identify cues for associative learning to be conscious. Indeed, they could learn associations between risky response outcomes and tiny fragments of the visual dynamic pattern formed by the mask/cue/mask flashing. However, post-conditioning debriefing questions might be informative in explaining why subjects could not discriminate between masked cues. Thus, when we explicitly asked subjects to state what the cues looked like, they reported in majority of cases that they had no idea. When the subjects were presented with the cues, now unmasked, they reported surprise at seeing the symbols while asserting that they had never seen them before. This suggests that during conditioning, subjects had no a priori representational knowledge to guide a visual search for cues hidden behind the masks. We believe that absence of an a priori representation is a crucial feature of our design, which, in addition to visual masking, prevented subjects from consciously seeing the cues.

Using this methodology, we show that pairing rewards and punishments can guide behavioral responses and even condition preferences for abstract cues that subjects could not consciously see. Note that if cues were visible, learning curves would have been optimized in one trial or two; hence we are not claiming that conscious awareness is unhelpful in supraliminal instrumental conditioning. However, in our subliminal conditioning task, conscious strategies (such as win-stay/lose-switch) might have been detrimental, which would explain why learning curves were limited well below the optimum.

We also identified brain circuitry associated with subliminal instrumental conditioning. The ventral striatum responded to subliminal cues and to visible outcomes in a manner that closely approximates Q-learning algorithm, expressing reward expected values and prediction errors, just as was reported in supraliminal instrumental conditioning studies (O'Doherty et al., 2004; Pagnoni et al., 2002; Pessiglione et al., 2006; Yacubian et al., 2006). Interestingly, there is no need for representing the statistical structure of the task in Q learning, which is an incremental procedure updating the expected values of chosen actions according to the subject's history of reward and punishment outcomes. This accords well with views that the striatum is a major player in implicit/procedural learning (Graybiel, 2005; Hikosaka et al., 1999; Packard and Knowlton, 2002) and with evidence that ventral striatum encodes reward-related information (Delgado, 2007; Knutson and Cooper, 2005; Pecina et al., 2006).

For the sake of simplicity, we have described ventral striatum activity as directly reflecting key outputs of Q-learning algorithm: Q value at the time of cue onset and prediction error at the time of outcome. There are nonetheless other variables in machine learning literature that would also correlate with ventral striatum activity and which could provide an alternative interpretational framework for our study. In particular, it is important to note that average Q values (over the reward, neutral, and punishment conditions) remain around zero during our conditioning paradigm. Hence, Q value is approximately equal to Q value minus average Q value, which can be seen as equivalent to a cue prediction error (actual Q value minus predicted Q value). Our data are therefore equally compatible with the notion that the ventral striatum encodes prediction errors at the time of both cue and outcome onsets. However, because prediction errors represent a function of Q values, the brain has to learn about Q values in order to signal prediction errors. Thus, whether we consider the ventral striatum as encoding a Q value or a prediction error does not alter our central conclusion: namely, the human brain can learn the reward value of subliminal cues, so as to later influence behavioral choices.

It is of interest that extrastriate visual cortex also reflected the reward value of subliminal cues, but not outcome-related prediction errors. Modulation of visual cortex activity by monetary incentives has already been reported in neuroimaging studies of supraliminal processes, such as visuomotor transformation, attentional control, and working memory (Krawczyk et al., 2007; Ramnani and Miall, 2003; Small et al., 2005). In our case, the modulation suggests that conditioning involves an interaction between the extrastriate cortex (which would discriminate cues according to their visual properties) and the ventral striatum (which would tag cues with affective values depending on reward prediction errors). However, we acknowledge that we do not as yet have a complete account of how the brain produces behavioral effects of subliminal conditioning. Notably, we failed to identify the brain regions mapping affective values onto motor commands, which would complete the circuit from visual cues to behavioral responses. Further experiments will be necessary to fill in these explanatory gaps.

More generally, our approach, combining perceptual masking and computational modeling, can be extended over the field of functional neuroimaging. Computational reinforcement learning

theory has proven useful to model both brain activity and behavioral choices in human and nonhuman primates (Daw and Doya, 2006; McClure et al., 2004; O'Doherty et al., 2007). Brain activity reflecting sophisticated computations are unlikely to be accessed by the conscious mind, which takes minutes to solve even simple equations. This brain activity would therefore represent unconscious processes, which we formally demonstrated here in the case of instrumental conditioning. Combining masking and modeling can, in principle, make more tractable the identification of basic neuronal mechanisms shared within other species, eliminating the use of reportable knowledge that might be unique to humans. It might also help assess the integrity of these same basic mechanisms in patients with neurological or psychiatric conditions, avoiding confounds generated by conscious strategic compensations.

EXPERIMENTAL PROCEDURES

Subjects

The study was approved by the National Hospital for Neurology and Neurosurgery and the Institute of Neurology joint Ethics Committee. Subjects were recruited via Gumtree website and screened for exclusion criteria: left handedness, age below 18 or above 39, regular taking of drug or medication, history of psychiatric or neurological illnesses and contra-indications to MRI scanning (pregnancy, claustrophobia, metallic implants). All subjects gave informed consent prior to taking part. We scanned 20 subjects: 11 males (mean age 26.8 ± 6.3 years) and 9 females (mean age 23.8 ± 3.3 years). Two more subjects were scanned but discarded from the analysis, because they eventually could describe parts of the visual cues, were above chance level in the perception task, and won unusually large amounts of money in the conditioning task. Subjects were told that they would play for real money, but at the end of the experiment their winnings were rounded up to a fixed amount.

Behavioral Task and Analysis

Subjects first read the instructions (see Supplemental Data) about the different tasks, which were later explained again step by step. Before scanning, subjects were trained to perform the conditioning task and the perception task on practice versions. In the scanner, they had to perform three sessions of the conditioning task, each containing 120 trials and lasting 13 min, and one session of the perception task, containing 120 trials and lasting about 7 min. The abstract cues were letters taken from the Agathodaimon font. The 12 cues shown in the scanner were randomly assigned to the four task sessions, each session hence employing 3 new cues. The same two masking patterns (see Figure 1), one displayed before and the other after the cue, were used in all task sessions. The sequence of display and the cue-outcome associations were also randomized for every subject.

The perceptual discrimination task was used to select the appropriate duration for cue display, which was then kept to either 33 or 50 ms for the entire experiment. In this task, subjects were flashed two masked cues, 3 s apart, displayed on the center of a computer screen, each following a fixation cross. They had to report whether or not they perceived any difference between the two visual stimulations. The response was given manually, by pressing one of two buttons assigned to "same" and "different" choices. The perceptual discrimination task was then employed as a control for awareness at the end of conditioning sessions. We checked with a chi-square test that in all included subjects performance was not significantly different from chance level (50% of correct responses). We also calculated d' measure, which is the difference between normalized rates of hits (correct "different" response) and false alarms (incorrect "different" responses). We ensured that this measure was not significantly different from zero, at group level, using one-tailed paired t test.

The instrumental conditioning task involved choosing between pressing or not pressing a button, in response to masked cues. After showing the fixation cross and the masked cue, the response interval was indicated on the

computer screen by a question mark. The interval was fixed to 3 s and the response was taken at the end: "Go" if the button was being pressed, "No" if the button was released. The response was written on the screen as soon as the delay had elapsed. Subjects were told that one response was safe (you do not win or lose anything) while the other was risky (you can win £1, lose £1, or get nothing). The risky response was assigned to Go for half of the subjects, and to NoGo for the other half, such that motor aspects were counterbalanced between reward and punishment conditions. Subjects were also told that the outcome of the risky response would depend on the cue that was displayed between the mask images. In fact, three cues were used, one was rewarding (+£1), one was punishing (−£1), and the last was neutral (£0). Because subjects were not informed about the associations, they could only learn them by observing the outcome, which was displayed at the end of the trial. This was either a circled coin image (meaning +£1), a barred coin image (meaning −£1), or a gray square (meaning £0).

Subjects were then debriefed about their visual perceptions and their response strategies. They reported responding either at chance, following their intuition, or following logical rules. None of them had the slightest idea of what the cues looked like. For the preference judgment task, the cues were then shown unmasked on a computer screen. The three cues used for a given session were displayed side by side, the position being randomized. Subjects were asked to rate them in order of preferences: 3 for the most liked, 2 for the intermediate, and 1 for the least one.

To assess for instrumental conditioning, we used one-tailed paired t tests comparing individual earnings with chance level (which is £0). Similarly, to assess for preference conditioning, we used one-tailed paired t tests comparing differential rating of winning and losing cues with chance level (which is 0).

Computational Model

We used a standard Q-learning algorithm (Sutton and Barto, 1998), which has been shown previously to offer a good account of instrumental choice in both humans and monkeys (Daw and Doya, 2006; McClure et al., 2004; O'Doherty et al., 2007). For each cue, the model estimates the expected value of the risky response, on the basis of individual sequences of choices and outcomes. This value, termed a Q value, is essentially the amount of reward expected from choosing the risky response given the contextual cue. These Q values were set at 0.1 before learning, and after every risky response the value of the cue was updated according to the Rescorla-Wagner rule: $Q(t+1) = Q(t) + \alpha * \delta(t)$. Following this rule, values are increased if the outcome is better than expected, and decreased in the opposite case. The prediction error was $\delta(t) = R(t) - Q(t)$, with $R(t)$ defined as the reinforcement obtained from choosing the risky response at trial t. In other words, the prediction error $\delta(t)$ is the difference between the expected outcome, i.e., $Q(t)$, and the actual outcome, i.e., $R(t)$. The reinforcement magnitude R was +1 and −1 for winning and losing £1, and 0 for neutral outcomes. Given the Q value, the associated probability of choosing the risky response was estimated by implementing the softmax rule: $P(t) = 1 / (1 + \exp(-Q(t)/\beta))$. This rule ensures that likelihood will be superior to 0.5 for positive values and inferior to 0.5 for negative values. The learning rate α concerns the amplitude of value changes from one trial to the next. The temperature β concerns the randomness of decision making. These two free parameters, α and β , were adjusted to maximize the probability (or likelihood) of the actual choices under the model. With the constraint that the parameters should be identical for reward and punishment cues we found: $\alpha = 0.1$ and $\beta = 0.9$. The model was then used to create statistical regressors corresponding to the Q values and prediction errors, for analysis of brain images.

Images Acquisition and Analysis

T₂-weighted echo planar images (EPI) were acquired with blood oxygen-level dependent (BOLD) contrast on a 3.0 Tesla magnetic resonance scanner. We employed a tilted plane acquisition sequence designed to optimize functional sensitivity in the orbitofrontal cortex and medial temporal lobes (Deichmann et al., 2003). To cover the whole brain with a short TR (1.95 s), we used the following parameters: 30 slices, 2 mm slice thickness, 2 mm interslice gap. T1-weighted structural images were also acquired, coregistered with the mean EPI, normalized to a standard T1 template, and averaged across subjects to allow group level anatomical localization. EPI images were analyzed in an event-related manner, within a general linear model, using the statistical para-

metric mapping software SPM5 (Wellcome Trust center for NeuroImaging, London, UK). The first 5 volumes of each session were discarded to allow for T1 equilibration effects. Preprocessing consisted of spatial realignment, normalization using the same transformation as structural images, and spatial smoothing using a Gaussian kernel with a full-width at half-maximum of 6 mm.

We used two different statistical linear regression models for our analyses. In both every trial was modeled as having two time points, corresponding to cue and outcome onsets. In the first model, two separate regressors were created for cues and outcomes, respectively modulated by the Q values and prediction errors computed by our optimized algorithm. In the second model, 12 separate regressors were created corresponding to the two time points (cues and outcomes) times the two conditioning phases (first and second half of each session) times the three conditions (reward, neutral, and punishment). In all cases, the regressors of interest were convolved with a canonical hemodynamic response function (HRF). To correct for motion artifact, subject-specific realignment parameters were modeled as covariates of no interest. Linear contrasts of regression coefficients were computed at the individual subject level and then taken to a group level random-effects analysis. At group level, we performed two statistical analyses: first a one-sample t test to find brain regions where regression coefficients were significant across subjects, and second a correlation with individual payoffs to find brain regions where regression coefficients increased with higher conditioning effect. A threshold of $p < 0.05$ after familywise error (FWE) correction for multiple comparisons was applied to avoid any a priori on brain localization. A more liberal threshold ($p < 0.001$, uncorrected) was also used to observe the extension of significant activations. To further illustrate activations, time courses were estimated by fitting a flexible basis set of finite impulse responses (FIRs), separated from the next by one scan (1.95 s). Both regression coefficients and time courses were then averaged across subjects, pooling together the voxels that passed the conservative threshold in statistical parametric maps (SPMs).

SUPPLEMENTAL DATA

The Supplemental Data include one figure and supplemental text and can be found with this article online at <http://www.neuron.org/cgi/content/full/59/4/561/DC1/>.

ACKNOWLEDGMENTS

This work was funded by the Wellcome Trust research program grants to C.D.F. and R.J.D. M.P. received a Young Researcher Prize from the Betten-court-Schuller Foundation, P.P. was supported by the Swedish Research Council, and J.D. is a beneficiary of the Marie Curie Intra-European research fellowship program. We wish to thank Helen Bates and Maël Lebreton, who independently replicated the behavioral findings outside the scanner.

Accepted: July 1, 2008

Published: August 27, 2008

REFERENCES

- Bayley, P.J., Frascino, J.C., and Squire, L.R. (2005). Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature* 436, 550–553.
- Berns, G.S., Cohen, J.D., and Mintun, M.A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science* 276, 1272–1275.
- Clark, R.E., and Squire, L.R. (1998). Classical conditioning and brain systems: the role of awareness. *Science* 280, 77–81.
- Cohen, N.J., Eichenbaum, H., Deacedo, B.S., and Corkin, S. (1985). Different memory systems underlying acquisition of procedural and declarative knowledge. *Ann. N Y Acad. Sci.* 444, 54–71.
- Daw, N.D., and Doya, K. (2006). The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 16, 199–204.

- Deichmann, R., Gottfried, J.A., Hutton, C., and Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19, 430–441.
- Delgado, M.R. (2007). Reward-related responses in the human striatum. *Ann. N Y Acad. Sci.* 1104, 70–88.
- Destrebecqz, A., and Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychon. Bull. Rev.* 8, 343–350.
- Graybiel, A.M. (2005). The basal ganglia: learning new tricks and loving it. *Curr. Opin. Neurobiol.* 15, 638–644.
- Hikosaka, O., Nakahara, H., Rand, M.K., Sakai, K., Lu, X., Nakamura, K., Miyachi, S., and Doya, K. (1999). Parallel neural networks for learning sequential procedures. *Trends Neurosci.* 22, 464–471.
- Knight, D.C., Nguyen, H.T., and Bandettini, P.A. (2003). Expression of conditional fear with and without awareness. *Proc. Natl. Acad. Sci. USA* 100, 15280–15283.
- Knutson, B., and Cooper, J.C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* 18, 411–417.
- Krawczyk, D.C., Gazzaley, A., and D'Esposito, M. (2007). Reward modulation of prefrontal and visual association cortex during an incentive working memory task. *Brain Res.* 1141, 168–177.
- Lagnado, D.A., Newell, B.R., Kahan, S., and Shanks, D.R. (2006). Insight and strategy in multiple-cue learning. *J. Exp. Psychol. Gen.* 135, 162–183.
- Lovibond, P.F., and Shanks, D.R. (2002). The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *J. Exp. Psychol. Anim. Behav. Process.* 28, 3–26.
- McClure, S.M., York, M.K., and Montague, P.R. (2004). The neural substrates of reward processing in humans: the modern role of fMRI. *Neuroscientist* 10, 260–268.
- Milner, B., Squire, L.R., and Kandel, E.R. (1998). Cognitive neuroscience and the study of memory. *Neuron* 20, 445–468.
- Morris, J.S., Öhman, A., and Dolan, R.J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature* 393, 467–470.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.
- O'Doherty, J.P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Ann. N Y Acad. Sci.* 1104, 35–53.
- Olsson, A., and Phelps, E.A. (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychol. Sci.* 15, 822–828.
- Packard, M.G., and Knowlton, B.J. (2002). Learning and memory functions of the Basal Ganglia. *Annu. Rev. Neurosci.* 25, 563–593.
- Pagnoni, G., Zink, C.F., Montague, P.R., and Berns, G.S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nat. Neurosci.* 5, 97–98.
- Pecina, S., Smith, K.S., and Berridge, K.C. (2006). Hedonic hot spots in the brain. *Neuroscientist* 12, 500–511.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., and Frith, C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.
- Poldrack, R.A., and Packard, M.G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia* 41, 245–251.
- Ramnani, N., and Miall, R.C. (2003). Instructed delay activity in the human prefrontal cortex is modulated by monetary reward expectation. *Cereb. Cortex* 13, 318–327.
- Seitz, A.R., and Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature* 422, 36.
- Small, D.M., Gitelman, D., Simmons, K., Bloise, S.M., Parrish, T., and Mesulam, M.M. (2005). Monetary incentives enhance processing in brain regions mediating top-down control of attention. *Cereb. Cortex* 15, 1855–1865.
- Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231.
- Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning (Cambridge, MA: MIT Press).
- Thorndike, E.L. (1911). Animal Intelligence: Experimental Studies (New York: Macmillan).
- Wilkinson, L., and Shanks, D.R. (2004). Intentional control and implicit sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 354–369.
- Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D.F., and Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* 26, 9530–9537.

Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions

Hilke Plassmann, John O'Doherty, and Antonio Rangel

California Institute of Technology, Division of Humanities and Social Sciences, MC 228-77, Pasadena, California 91125

An essential component of every economic transaction is a willingness-to-pay (WTP) computation in which buyers calculate the maximum amount of financial resources that they are willing to give up in exchange for the object being sold. Despite its pervasiveness, little is known about how the brain makes this computation. We investigated the neural basis of the WTP computation by scanning hungry subjects' brains using functional magnetic resonance imaging while they placed real bids for the right to eat different foods. We found that activity in the medial orbitofrontal cortex and in the dorsolateral prefrontal cortex encodes subjects' WTP for the items. Our results support the hypothesis that the medial orbitofrontal cortex encodes the value of goals in decision making.

Key words: decision making; reward; neuroeconomics; orbitofrontal cortex; fMRI; buying

Introduction

An essential component of every marketplace transaction is a willingness-to-pay (WTP) computation in which buyers calculate the maximum amount of resources that they are willing to give up in exchange for the object being sold. The WTP computation is used to evaluate whether a proposed trade is beneficial (e.g., when the WTP exceeds the price at which the item is being offered) or to decide how much to bid for an item (e.g., when competing with other individuals at an auction). To make good trades, individuals must be able to assign a WTP to an item that is commensurate to the benefits that it will generate. Otherwise they would end up purchasing items for a price that exceeds their worth to them. Despite its pervasiveness and importance for economic well being, little is known about how the brain performs the WTP computation in everyday transactions, or about how its ability to do so is affected by diseases such as addiction or obsessive compulsive disorders. This makes understanding how and where the brain makes these computations one of the most important open questions in the nascent field of neuroeconomics (Glimcher and Rustichini, 2004; Camerer et al., 2005).

Based on the results of several previous studies, we hypothesized that activity in the medial orbitofrontal cortex (mOFC) encodes WTP. Monkey electrophysiology studies of binary choice have found that activity in the OFC encodes the value of the available actions (Wallis and Miller, 2003; Padoa-Schioppa and Assad, 2006). Using functional magnetic resonance imaging (fMRI), Paulus and Frank (2003) found greater medial OFC activation during hypothetical choices than during a visual discrim-

ination task (see also Arana et al., 2003). Erk et al. (2002) found that mOFC activity during a hypothetical liking rating task increased with the reported attractiveness of the stimuli. Finally, a series of stimulus–outcome learning studies (in which no decisions were made) have shown that the mOFC maintains a representation of the expected reward associated with particular cues (Rolls, 1996; Schoenbaum et al., 1998; Tremblay and Schultz, 1999; Roesch and Olson, 2004). Although all of these studies suggest that the medial OFC plays a critical role in the evaluation of choices, none of them have established that activity in the medial OFC correlates with the economic computation of WTP.

We investigated the neural basis of the WTP computation by scanning hungry subjects' brains using fMRI while they placed bids for the right to eat different foods in a Becker–DeGroot–Marshak auction (Becker et al., 1964). The results described below confirmed our hypothesis: we found that activity in the right medial OFC encodes subjects' WTP for items.

Materials and Methods

Subjects. Nineteen normal-weight subjects participated in the experiment (16 males, mean age, 25.45; age range, 18–46). One additional subject participated in the experiment but was excluded from the analysis because she did not understand the instructions. All subjects were right-handed, healthy, had normal or corrected-to-normal vision, had no history of psychiatric diagnoses, neurological or metabolic illnesses, and were not taking medications that interfere with the performance of fMRI. All subjects had no history of eating disorders and were screened for liking and at least occasionally eating the types of foods that we used. Subjects were told that the goal of the experiment was to study food preferences and gave written consent before participating. Caltech's institutional The review board of the California Institute of Technology (Pasadena, CA) approved the study.

Stimuli. Subjects bid on 50 different sweet and salty junk foods (e.g., chips and candy bars). We selected the foods based on pilot data to satisfy several characteristics. First, we wanted items to be highly familiar and to be sold in local convenience stores, to remove uncertainty considerations from the WTP computation as much as possible. Familiarity data collected at the end of the experiment shows that we were successful [famil-

Received May 9, 2007; revised July 2, 2007; accepted July 24, 2007.

This work was supported by the Moore Foundation, National Science Foundation Grant SES-0134618, and German Academic Exchange Service Grant DAAD D/05/47698. We thank Vivian Valentin, Jan Glaescher, Alan Hampton, and Axel Linder for their help with this work.

Correspondence should be addressed to Antonio Rangel, California Institute of Technology, Division of Humanities and Social Sciences, MC 228-77, Pasadena, CA 91125. E-mail: rangel@hss.caltech.edu.

DOI:10.1523/JNEUROSCI.2131-07.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/279984-05\$15.00/0

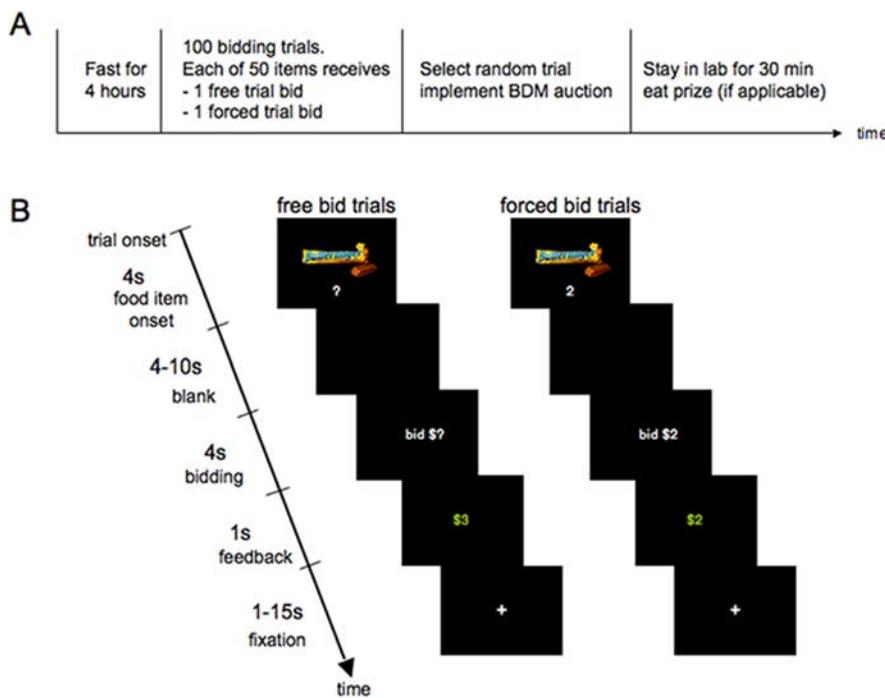


Figure 1. Experimental design. *A*, Timeline of the experiment. *B*, Time course for free bid and forced bid trials. Free and forced bid trials were identical except that in forced bid trials visual presentation of the food items was paired with the forced bid amount. In addition, the forced bid amount was repeated during the bidding cue. Food items, trial type, and forced bid amounts were fully randomized within subjects.

arity scores: mean, 3.97; SD, 1.34; scale, 1 (not familiar) to 5 (very familiar)]. Second, we wanted items to be positive for the subjects (in the sense that their WTP for them is greater or equal than zero). The foods were presented to the subjects using high-resolution color pictures (72 dpi). The stimulus presentation and response recording was controlled by E-prime (Psychology Software Tools, Pittsburgh, PA). The visual stimuli were presented using video goggles.

Task. Figure 1 describes the time structure of the experiment. Subjects were instructed not to eat for 4 h before the experiment, which increased the value that they placed on the foods. They were also instructed that they would have to remain in the lab for 30 min at the conclusion of the experiment, and that the only thing that they will be able to eat is whatever food they purchased from us during the task. In addition to a \$35 participation fee, each subject received three \$1 bills in “spending money” to purchase food from us. Whatever money they did not spend was theirs to keep.

Subjects placed bids for the right to eat a snack at the end of the experiment in 100 different bidding trials. In each trial they were allowed to bid \$0, \$1, \$2, or \$3 for each food item. At the end of the experiment, one of those trials was randomly selected, by drawing a ball from an urn, and only the outcome of that trial was implemented. As a result, subjects did not have to worry about spreading their \$3 dollar budget over the different items and they could treat each trial as if it were the only decision that counted. Objects were sold using a Becker–DeGroot–Marschak auction. The rules of the auction are as follows. Let b denote the bid made by the subject for a particular item. After the bid is made a random number n is drawn from a known distribution (in our case, \$0, \$1, \$2, and \$3 were chosen with equal probability). If $b \geq n$, the subject got the item and paid a price equal to n . In contrast, if $b < n$, the subject did not get the object but also did not have to pay anything.

We used this auction institution as our model of market transactions in the laboratory because it has three very useful properties. First, it is characterized by a simple set of rules. Second, the optimal strategy for a buyer is to bid exactly her WTP for the item being sold (Becker et al., 1964). The intuition for why this is the case is as follows. There is no incentive to bid less than the WTP because the price paid is determined by the random number n and, thus, the bids do not affect the price paid.

There is also no incentive to increase the bid above the WTP because this may lead to a situation in which the subject gets the item but ends up paying a price larger than his WTP (e.g., consider the case $\text{WTP} = \$1$, $b = \$3$, and $n = \$2$). The fact that bidding the WTP is the optimal strategy was explained and emphasized extensively during the instruction and training period. We performed an extensive amount of pilot work to find a set of instructions that led to 100% reported compliance with the best strategy. The instructions, included in supplemental material (available at www.jneurosci.org), emphasized that the subject’s best strategy is to look at the item, ask themselves how much is worth, and simply bid that amount. Third, because individuals always bid their exact WTP, we got a measure of the WTP computed by the brain for every bidder and item at the time of decision making, which we could then compare with the blood oxygenation level-dependent (BOLD) measure of neural activity.

To keep the task simple, subjects were only allowed to bid discrete amounts for the items (\$0, \$1, \$2, or \$3). A consequence of this is that the bids are only approximations of the true WTP computed by subjects. For example, when the true WTP is \$2.3, our measure is \$2. Similarly, subjects with a WTP larger than \$3 enter a bid of \$3. However, the bids are a monotonic function of the true WTP and highly correlated with it.

We used two different kinds of trials: free-bid trials and forced-bid trials. Each of the 50 items was shown twice, once in a bid trial and once in a forced trial. These trials were fully randomized within and across subjects. Both types of trials had an equal probability of being selected to be the trial that counted. The timing for each type of trial is shown in Figure 1*B*. The only difference between the two types of trials is that whereas subjects were free to select the amount of their bid in the free trials, they were told how much to bid in the forced trials. The forced bids were drawn uniformly and independently from \$0, \$1, \$2, or \$3 on each trial. The set of rules described above applied to both trials. Note that subjects needed to make a willingness-to-pay computation in free trials to decide how much to bid, but they did not need to do so in forced trials.

After receiving the instructions, subjects were trained on using the response boxes with their right hand and on the bidding procedure. To avoid activation artifacts caused by the assignment of buttons to bid amounts, the assignment was counterbalanced across subjects.

The existence of two types of bidding trials is a novel and essential component of the experimental design. A difficulty in searching for the neural basis of the WTP computation is that, when the brain is exposed to a picture of a food item, it might calculate other variables that are correlated with WTP. For example, the brain may simulate the anticipated taste of the food, or it may assess its caloric content. If this issue is not properly addressed, one could erroneously attribute WTP computations to areas that are calculating different albeit correlated variables. The presence of free and forced trials provides a solution to the problem. The only difference between both types of trials is that the subject needs to perform a WTP computation in the free trials, because she needs to decide how much to bid, but not in the forced trials, because she is told what her bid should be. Every other computation, such as the anticipated taste of the food, should be performed equally in both types of trials. As a result, we can conclude that a brain area encodes the WTP computation whenever its activity increases with the WTP in the free trials, but not in the forced trials.

fMRI data acquisition. The functional imaging was conducted using a Siemens (Erlangen, Germany) 3.0 Tesla Trio MRI scanner to acquire gradient echo T2*-weighted echoplanar (EPI) images with BOLD con-

trast. To optimize functional sensitivity in the OFC, we used a tilted acquisition in an oblique orientation of 30° to the anterior commissure-posterior commissure line (Deichmann et al., 2003). In addition, we used an eight-channel phased array coil which yields a 40% signal increase in signal in the medial OFC over a standard head coil. Each volume comprised 32 axial slices. A total of 1100 volumes (two sessions, ~18 min each) were collected during the experiment in an interleaved-ascending manner. The imaging parameters were as follows: echo time, 30 ms; field of view, 192 mm; in-plane resolution and slice thickness, 3 mm; repetition time, 2 s. Whole-brain high-resolution T1-weighted structural scans ($1 \times 1 \times 1$ mm) were acquired from the 19 subjects and coregistered with their mean EPI images and averaged together to permit anatomical localization of the functional activations at the group level. Image analysis was performed using SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Temporal normalization was applied to the scans with a time of acquisition of 1.9375 referenced to the last volume. To correct for subject motion, the images were realigned to the last volume, spatially normalized to a standard T2* template with a resampled voxel size of 3 mm, and spatially smoothed using a Gaussian kernel with a full width at half maximum of 8 mm. Intensity normalization and high-pass temporal filtering (using a filter width of 128 s) were also applied to the data.

fMRI data analysis. The data analysis proceeded in three steps. First, we estimated a general linear model with AR(1) and the following regressors that capture the main events in our experiment: free bid and picture presentation (R1), free bid and response (R2), forced bid and picture presentation (R3), forced bid and response (R4), missed bid trial and picture presentation (R5), and missed bid trial and response (R6). The regressors that capture the presentation of the food pictures were modeled using 4 s box-car functions. The regressors for the bid responses were modeled using stick functions.

To take advantage of the parametric nature of our design, the general linear model also included the following parametric modulators: free bid and picture presentation modulated by bid (M1), free bid and picture presentation modulated by surplus₊ (M2), free bid and response modulated by bid (M3), free bid and response modulated by surplus₊ (M4), forced bid and picture presentation modulated by bid (M5), forced bid and item presentation modulated by surplus₊ (M6), forced bid and item presentation modulated by surplus₋ (M7), forced bid and response modulated by bid (M8), forced bid and response modulated by surplus₊ (M9), and forced bid and response modulated by surplus₋ (M10). The parametric modulators are defined as follows. “Bid” equals the amount bid for the item sold in that trial during the corresponding free trial and, thus, is a measure of the subject’s WTP for the item being shown. “Surplus” is a variable that measures the expected “profit” from the trial given the bid that was placed (and conditional on the trial being selected to be the one that counts). For example, suppose that a subject’s true value is \$2 and that he bids \$2. Then the surplus equals $0.25 \times \$2 + 0.25 \times \$1 + 0.50 \times \$0 = \0.75 , where the first term measures the probability that the random number in the auction is 0 times the profit made in that case, the second number measure the probability that the random number is 1 times the profit made in that case, and so on. Surplus₊ = max[0, surplus] and Surplus₋ = min[surplus, 0]. (Note that for the free trials, the surplus variable is always non-negative, whereas for the forced trials it can be positive or negative.) We orthogonalized the modulators for each of the main regressors (M1 and M2, M3 and M4, M5 to M7, and M8 to M10). Each of the regressors was convolved with a canonical hemodynamic response function. We also included a constant term and six motion parameters as regressors of no interest. Note that the inclusion of the surplus modulator is important to avoid confounding areas that code for WTP with areas that code for economic surplus.

Second, we calculated the following first-level single-subject contrasts: (1) free-bid trials when exposed to item modulated by bid (regressor M1), (2) forced-bid trials when exposed to item modulated by bid (regressor M5), and (3) free- minus forced-bid trials when exposed to an item modulated by bid (regressors M1 minus M5).

Third, we calculated second-level group contrasts using a one-sample *t* test. The figures shown below are constructed using these second-level contrasts at a threshold of $p < 0.001$, uncorrected, and a minimum

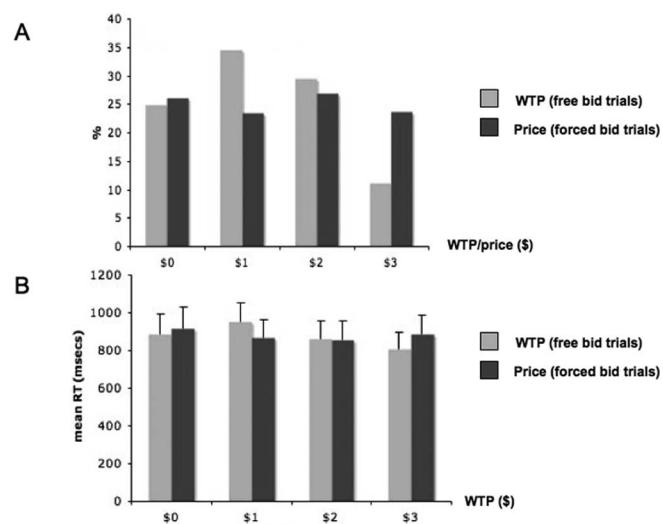


Figure 2. Behavioral results. *A*, Distribution of bids in free and forced trials. *B*, Reaction times for free and forced trials as a function of bid. Error bars denote SEs.

cluster size of 10. Anatomical localizations were then performed by overlaying the *t* maps on a normalized structural image averaged across subjects, and with reference to an anatomical atlas (Duvernoy, 1999).

Results

Behavioral

Figure 2*A* shows the distribution of bids during free- and forced-bid trials. The average free bid was \$1.4 (SD, 0.27) and over 75% of the free bids were greater than zero. The bid amounts for the forced bid trials were randomly drawn from a uniform distribution on \$0, \$1, \$2, and \$3. Although there is substantial variability on value that subjects place on particular items, the average WTP was significantly greater than zero ($p < 0.001$), which suggests that most items were rewarding for most subjects.

Figure 2*B* shows the distribution of bidding reaction times for free and forced trials. The reaction times were entered in a two-way repeated-measures ANOVA with two factors: bid amount (\$0, \$1, \$2, or \$3) and trial type (free- or forced-bid trial). The analysis revealed no significant main effects or interactions.

Neuroimaging results

Identifying the neural correlates of WTP in free trials

We performed a whole-brain analysis to identify areas that correlated with WTP in the free trials at the time of evaluation (i.e., when the food picture is displayed). This contrast is interesting because it identifies areas that might encode for WTP. Our hypothesis was that activity in the OFC would be positively correlated with WTP. The hypothesis was supported by the data: activity in the medial OFC ($x = 6, y = 30, z = -17; p < 0.001$, uncorrected) was correlated with WTP. Other areas identified by this contrast were the dorsal anterior cingulate cortex ($x = -4, y = 34, z = -20; p < 0.001$, uncorrected), and the dorsolateral prefrontal cortex (DLPFC; $x = 44, y = 44, z = 12; p < 0.001$, uncorrected).

Identifying the neural correlates of value during the forced-bid trials

Our experimental design is based on the idea that the brain computes a WTP during free-bid trials, but not during forced-bid trials. To test this hypothesis, we performed a whole-brain analysis to identify areas correlated with WTP in the forced-bid trials at the time of evaluation (i.e., when the food picture is displayed).

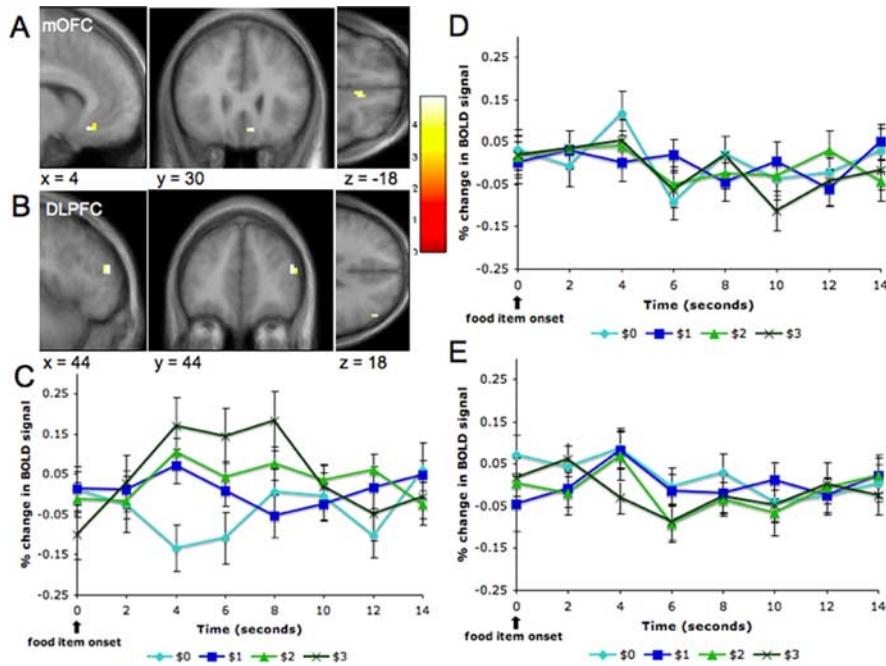


Figure 3. Neural correlates of WTP. **A, B**, Activity in the medial OFC and the DLPFC was positively correlated with WTP at the time of evaluation in the free trials more than in the forced trials. Activation maps shown at a threshold of $p < 0.001$ uncorrected and 10 voxel clusters. **C**, Averaged time courses for the medial OFC voxels during free trials as a function of WTP (error bars denote SEs). **D**, Averaged time courses for the medial OFC voxels during forced bid trials as a function of WTP. **E**, Averaged time courses for the medial OFC voxels during forced bid trials as a function of the forced bid. A comparison of the time courses shows that the medial OFC encodes WTP in free trials, but not in forced trials, and that it does not encode the forced bid amounts.

No areas showed the desired correlation at a level of $p < 0.001$ (uncorrected) and an extent threshold of 10 voxels.

Another possibility is that during the forced trials the brain encodes either the size of the forced bid, or the disagreement between the forced bid and the WTP. We tested for both possibilities using minor variants of the general linear model described above and found no regions of correlation at a level of $p < 0.001$ (uncorrected) and an extent threshold of 10 voxels (for details, see the supplemental material, available at www.jneurosci.org).

Test of the main hypothesis

As discussed above, a limitation of the previous contrast is that it identifies areas with activity that is correlated with WTP, but also areas that encode for variables that are correlated with it, such as anticipatory taste. To address this potential confound, we looked for areas that (1) showed increasing activation with WTP in the free trials and (2) were significantly more activated in the free trials than in the forced trials. As predicted, we found that the right mOFC satisfied these conditions ($x = 4, y = 30, z = -18; p < 0.001$, uncorrected). Unexpectedly, we also found that right DLPFC satisfied them ($x = 44, y = 44, z = 18; p < 0.001$, uncorrected). Figure 3, *A* and *B*, describes the results of this contrast.

We also extracted trial averaged time-course data from peak voxels in the mOFC for each subject, which were then averaged across subjects (Fig. 3*C–E*). The time courses show that activity in this area during free trials showed an increase in activation that was correlated with the subjects' bids. The time courses also show that activity during forced trials did not discriminate the subjects' WTP for the items (as measured by their bids for that item during free trials) or the magnitude of the forced bids. We thus concluded that activity in the right medial OFC and dorsolateral prefrontal cortex encode for WTP in everyday economic decisions.

Discussion

In this study, we provide evidence that mOFC encodes subjects' WTP during simple economic transactions. Critically, we used a parametric experimental design that allowed us to identify areas that encode for WTP, as opposed to areas that are active during economic choice, but that do not correlate with WTP (Blair et al., 2006; Arana et al., 2003; Paulus and Frank, 2003).

Our findings are consistent with data from human lesion studies showing that lesions to the ventromedial prefrontal cortex impair the ability to make consistent pairwise choices (Fellows and Farah, 2007). Our findings are also consistent with monkey electrophysiology studies of simple choice behavior. For example, several studies have found that OFC neurons encode for the decision or incentive value of the stimuli at the time of decision making (Wallis and Miller, 2003; Roesch and Olson, 2004; Padoa-Schioppa and Assad, 2006). In addition to providing cross-species and cross-modality validation, this experiment shows that the OFC plays a central role in the encoding of decision values in decision-making tasks that are significantly more abstract and complex than those that can be studied in monkey

experiments. As a result, we speculate that the mOFC might encode for the decision value of choices in a wide class of economic settings. However, because this study and much of the previous literature has focused on the valuation of primary appetitive rewards, such as desirable foods, additional work is needed to investigate whether the mOFC area also encodes for the value of nonprimary rewards, such as a book or a DVD, and of negative or undesirable items, such as electric shocks of different magnitudes.

An open question in behavioral neuroscience is which parts of the OFC play a role in learning the encoding of stimulus-outcome associations (Tremblay and Schultz, 1999; Rolls, 2000) and which parts are involved in guiding decisions by encoding the value of alternative goals (Arana et al., 2003; Schoenbaum and Roesch, 2005; Feierstein et al., 2006; Padoa-Schioppa and Assad, 2006)? Our results suggest that the mOFC plays a critical role in goal directed behavior by encoding economic value.

We also found that activity in the DLPFC was correlated with the subject's WTP. This is consistent with previous monkey neurophysiology studies that have found neurons in the DLPFC that encode aspects of the decision or the incentive value of stimuli (Watanabe, 1996; Wallis and Miller, 2003). This raises an important open question in the neuroeconomics literature: what are the relative contributions of the DLPFC and the mOFC to the valuation of stimuli during economic decision making? The neuroanatomy of these two regions suggests a potential explanation. The OFC receives inputs from multiple sensory areas, which are likely to be used in valuation, whereas the DLPFC does not (Price, 2006). In contrast, the DLPFC is heavily connected with motor output areas, whereas the OFC is not connected with these areas directly (Petrides and Pandya, 1999). Finally, the DLPFC and the

OFC are interconnected (Petrides and Pandya, 1999). This pattern of connectivity suggests that economic values might be first computed in the mOFC and then passed to the DLPFC to influence motor commands. This pathway is not unique as the OFC might also be able to affect actions through its connections to the striatum, which, in turn, is also heavily connected to the motor system (Yeterian and Pandya, 1991). This conjecture is consistent with our data, but our experimental design does not allow us to reach this directly.

It is interesting to compare our findings with those of Knutson et al. (2007), who study which brain areas are involved in making economic purchase decisions for unfamiliar items sold at exogenously given prices. Their task proceeds in three steps: (1) subjects are shown a picture of the item that is for sale, (2) the sale price is added, and (3) subjects decide whether to purchase the item or not. Knutson et al. (2007) obtained a measure of the value that the subjects placed on the items, albeit not a WTP, and looked for neural correlates of this value at the time of picture presentation (step 1). They found a correlation with nucleus accumbens (NAcc) activity, but not with OFC activity. In contrast, in our study activity in the medial OFC, but not in the NAcc, encoded the subject's WTP for items. Note that the difference in results is not attributable to an inability to image the NAcc because we get strong striatal activation in other contrasts of interest (supplemental Tables 3, 4, available at www.jneurosci.org as supplemental material). Instead, it is likely that there are subtle but important differences between the computations that the brain makes in the two tasks, and that the mOFC and NAcc play a differential role in such computations. For example, in the Knutson et al. (2007) task, the value of purchasing an item is the WTP minus the price and the information needed to compute this "net value" is revealed over time. In contrast, in our experiment all of the information needed to compute the WTP is revealed at the beginning of the trials. As a result, anticipatory reward signals, which are known to be correlated with NAcc activity (Knutson et al., 2001), might be computed in the Knutson et al. (2007) task, but not in the current study. Additional experiments are needed to systematically explore the differences between the computations made in the two experiments.

Part of the research agenda in neuroeconomics is to understand how the brain evaluates potential goals and outcomes at the time of decision making, and how other cognitive, emotional, and visceral processes affect the computation of economic value. A first step in this research agenda is to understand what are the brain structures responsible for the computation of value in simple everyday choices. Our results suggest that the medial OFC is a place where a variety of variables computed in other brain regions are integrated into a single representation of value. If this hypothesis is correct, other brain processes may be able to influence decision making by modulating activity in the medial OFC.

References

- Arana FS, Parkinson JA, Hinton E, Holland AJ, Owen AM, Roberts AC (2003) Dissociable contributions of the human amygdala and orbito-frontal cortex to incentive motivation and goal selection. *J Neurosci* 23:9632–9638.
- Becker G, DeGroot M, Marschak J (1964) Measuring utility by a single-response sequential method. *Behav Sci* 9:226–232.
- Blair K, Marsh AA, Morton J, Vytlilingam M, Jones M, Mondillo K, Pine DC, Drevets WC, Blair JR (2006) Choosing the lesser of two evils, the better of two goods: specifying the roles of ventromedial prefrontal cortex and dorsal anterior cingulate in object choice. *J Neurosci* 26:11379–11386.
- Camerer C, Loewenstein G, Prelec D (2005) Neuroeconomics: how neuroscience can inform economics. *J of Economic Literature* 43:9–64.
- Deichmann R, Gottfried JA, Hutton C, Turner R (2003) Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage* 19:430–441.
- Duvernoy HM (1999) The human brain: surface, three-dimensional sectional anatomy with MRI, and blood supply. Berlin: Springer.
- Erk S, Spitzer M, Wunderlich A, Galley L, Walter H (2002) Cultural objects modulate reward circuitry. *NeuroReport* 13:2499–2503.
- Feierstein CE, Quirk MC, Uchida N, Sosulski DL, Mainen ZF (2006) Representation of spatial goals in rat orbitofrontal cortex. *Neuron* 51:495–507.
- Fellows LK, Farah MJ (2007) The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex*, in press.
- Glimcher PW, Rustichini A (2004) Neuroeconomics: the concilience of brain and decision. *Science* 306:4447–4452.
- Knutson B, Adams CM, Fong GW, Hommer D (2001) Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci* 21:RC159.
- Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G (2007) Neural predictors of purchases. *Neuron* 53:147–156.
- Padoa-Schioppa C, Assad JA (2006) Neurons in the orbitofrontal cortex encode economic value. *Nature* 441:223–226.
- Paulus MP, Frank LR (2003) Ventromedial prefrontal cortex activation is critical for preference judgments. *NeuroReport* 14:1311–1315.
- Petrides M, Pandya DN (1999) Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur J Neurosci* 11:1011–1036.
- Price JL (2006) Connections of orbital cortex. In: the orbitofrontal cortex (Zald DH, Rauch SL, eds), pp 39–55. Oxford: Oxford UP.
- Roesch MR, Olson CR (2004) Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304:307–310.
- Rolls ET (1996) The orbitofrontal cortex. *Philos Trans R Soc Lond B Biol Sci* 351:1433–1443, discussion 1443–1434.
- Rolls ET (2000) The orbitofrontal cortex and reward. *Cereb Cortex* 10:284–294.
- Schoenbaum G, Roesch M (2005) Orbitofrontal cortex, associative learning, and expectancies. *Neuron* 47:633–636.
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1:155–159.
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.
- Wallis JD, Miller EK (2003) Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18:2069–2081.
- Watanabe M (1996) Reward expectancy in primate prefrontal neurons. *Nature* 382:629–632.
- Yeterian EH, Pandya DN (1991) Prefrontostriatal connections in relation to cortical architectonic organization in rhesus monkeys. *J Comp Neurol* 312:43–67.

The Architecture of Reward Value Coding in the Human Orbitofrontal Cortex

Guillaume Sescousse,^{1,2} Jérôme Redouté,^{1,2,3} and Jean-Claude Dreher^{1,2}

¹Center for Cognitive Neuroscience, Reward and Decision-Making Group, Centre National de la Recherche Scientifique, Unité Mixte de Recherche 5229, 69675 Bron, France, ²Université Lyon 1, 69003 Lyon, France, and ³CERMEP—Imagerie du Vivant, 69003 Lyon, France

To ensure their survival, animals exhibit a number of reward-directed behaviors, such as foraging for food or searching for mates. This suggests that a core set of brain regions may be shared by many species to process different types of rewards. Conversely, many new brain areas have emerged over the course of evolution, suggesting potential specialization of specific brain regions in the processing of more recent rewards such as money. Here, using functional magnetic resonance imaging in humans, we identified the common and distinct brain systems processing the value of erotic stimuli and monetary gains. First, we provide evidence that a set of neural structures, including the ventral striatum, anterior insula, anterior cingulate cortex, and midbrain, encodes the subjective value of rewards regardless of their type, consistent with a general hedonic representation. More importantly, our results reveal reward-specific representations in the orbitofrontal cortex (OFC): whereas the anterior lateral OFC, a phylogenetically recent structure, processes monetary gains, the posterior lateral OFC, phylogenetically and ontogenetically older, processes more basic erotic stimuli. This dissociation between OFC representations of primary and secondary rewards parallels current views on lateral prefrontal cortex organization in cognitive control, suggesting an increasing trend in complexity along a postero-anterior axis according to more abstract representations. Together, our results support a modular view of reward value coding in the brain and propose that a unifying principle of postero-anterior organization can be applied to the OFC.

Introduction

A basic concern about the functional organization of the prefrontal cortex is to delineate the functional divisions of the orbitofrontal cortex (OFC). A number of lesion, electrophysiological, and neuroimaging studies indicate a general role for the OFC in encoding the value assigned to different goods in both human (O'Doherty et al., 2003a; Plassmann et al., 2007; Chib et al., 2009; FitzGerald et al., 2009) and nonhuman primates (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2008). An important remaining issue, which is key to our understanding of the functional organization of the OFC, is to determine whether distinct parts of the OFC encode rewards of different nature. The anterior and posterior parts of the OFC are considered to belong to two distinct cytoarchitectonic trends. The anterior part of the OFC, especially well developed in humans and characterized by a granular cell layer, is thought to be phylogenetically and ontogenetically more recent than the posterior and medial parts, which consist of agranular and dysgranular cortices (Ongür and Price, 2000; Wise, 2008). Although never tested empirically, one fundamental hypothesis, based on this increasing trend in complexity

along a postero-anterior axis, is that the anterior part of the OFC would process secondary rewards, whereas the posterior part would process primary rewards (Kringelbach and Rolls, 2004).

Here, we directly tested this hypothesis with functional magnetic resonance imaging (fMRI) by comparing the brain responses to two experienced rewards: money and erotic pictures. These two rewards present significant evolutionary differences likely to be reflected at the cerebral level: whereas money is a secondary reward that appeared recently in human history and whose abstract value needs to be learned by association with primary reinforcers, erotic stimuli can be considered as primary rewards because they have an innate value and satisfy biological needs. We therefore hypothesized that monetary gains would recruit anterior OFC regions and erotic pictures would engage more posterior OFC regions. Despite their critical sociobiological importance, erotic stimuli have never been studied as reinforcers but rather as arousing stimuli in passive viewing paradigms focusing on sexual function (Redouté et al., 2000; Ponseti et al., 2006). However, erotic stimuli are clearly rewarding (Hamann et al., 2004), probably because sexual attractiveness, which may have evolved to enhance reproductive fitness, is an important cue for mate choice (Thornhill and Gangestad, 1999; Rhodes, 2006).

In addition to specialized OFC regions processing different types of rewards, we hypothesized the existence of common brain structures supporting general hedonic representations independent of reward type. To evaluate and compare the relative value of different rewards on a unique scale, it has been proposed that the brain may use a “common neural currency,” likely to be implemented in integrative reward regions such as the ventral striatum

Received July 7, 2010; revised July 19, 2010; accepted Aug. 9, 2010.

This work was funded by a Marie Curie International Reintegration Grant (J.-C.D.) and a Fyssen Foundation Grant (J.-C.D.). G.S. was funded by a scholarship from the French Ministry of Research and the Medical Research Foundation. We thank the staff of CERMEP—Imagerie du Vivant for helpful assistance with data collection.

Correspondence should be addressed to either Guillaume Sescousse or Jean-Claude Dreher, Centre de Neurosciences Cognitive, CNRS UMR 5229, 67 Boulevard Pinel, 69675 Bron Cedex, France, E-mail: gsescousse@isc.cnrs.fr or dreher@isc.cnrs.fr.

DOI:10.1523/JNEUROSCI.3501-10.2010

Copyright © 2010 the authors 0270-6474/10/3013095-10\$15.00/0

and ventromedial prefrontal cortex (vmPFC) (Montague and Berns, 2002; Sugrue et al., 2005; Rangel et al., 2008; Dreher, 2009). Consistent with this claim, a number of studies have shown that goal/decision values, reflecting the anticipated rewarding properties of a stimulus at the time of choice, were encoded in these regions regardless of reward type (McClure et al., 2007; Hare et al., 2008; Chib et al., 2009; Peters and Buchel, 2010). However, very few studies have investigated whether shared cerebral substrates are engaged when actually experiencing different types of rewards, i.e., during the computation of outcome/hedonic values (Izuma et al., 2008; Smith et al., 2010). Such a common representation of outcome value may be particularly useful to support the trading ability demonstrated by primates. For example, humans are willing to sacrifice money to view attractive faces (Hayden et al., 2007), and similarly male monkeys exchange meat for sex or sacrifice fluid for the opportunity to view female perinea (Deaner et al., 2005; Gomes and Boesch, 2009).

Building on these considerations, we designed a reward paradigm aiming to (1) determine whether the OFC is functionally divided depending on reward type and (2) identify common brain structures processing the experienced value of both monetary and erotic rewards. To further characterize the role of these brain regions in reward processing, we also manipulated reward intensity and probability and collected hedonic ratings after reward outcomes inside the scanner. Because erotic pictures are not quantifiable like money, it is unclear whether changes in the intensity of erotic pictures would affect the same brain regions as those responding to changes in monetary amounts. Likewise, it is unknown whether the concept of prediction error, reflecting the discrepancy between expected and actual rewards and primarily used with quantifiable rewards (such as money and juice), can be extended to erotic stimuli.

Materials and Methods

Participants

Eighteen right-handed volunteers (mean \pm SD age, 24 \pm 3.3 years) with no history of neurological or psychiatric disorders participated in this study. All of them were heterosexual males, because men are generally more responsive to visual sexual stimuli than women (Hamann et al., 2004) and to avoid the potential influence of the menstrual cycle known to have an effect on reward processing in women (Caldú and Dreher, 2007; Dreher et al., 2007). All subjects gave written informed consent to be part of the experiment, which was approved by the local ethics committee (Centre Léon Bérard, Lyon, France).

Motivation, which was a crucial element of our study, was closely controlled. First, sexual arousability was assessed at the time of screening through specific questionnaires, namely the Brief Sexual Function questionnaire (Reynolds et al., 1988) and the Sexual Arousal Inventory (Hoon and Chambliss, 1998). Of an initial pool of 22 subjects, two of them were excluded because they scored too low on the Sexual Arousal Inventory (mean score for all subjects, 91.1 \pm 14.6; scores of excluded subjects, 54 and 64). To further ensure that all participants would be in a similar state of motivation to see erotic stimuli, we asked them to avoid any sexual contact during a period of 24 h before the scanning session. Second, we sought to enhance the motivation for money by telling the subjects that the financial compensation for their participation would be calculated based on their winnings during the task. We also excluded two subjects presenting symptoms of depression as assessed by the 13-item version of the Beck Depression Inventory (Beck and Beck, 1972) (mean score for all subjects, 1.9 \pm 2.7; scores of excluded subjects, 6 and 10).

Task

Our protocol was inspired from the typical design of incentive delay tasks (Knutson et al., 2005; Abler et al., 2006) but included several modifica-

tions related to our questions. Experimental trials were divided into two phases: reward anticipation and outcome. During reward anticipation, a cue was presented, followed by a delay period and a discrimination task (Fig. 1). The cue carried three types of information regarding the upcoming reward: the red portion of a pie chart in the background indicated its probability (25, 50, or 75%), and the pictogram in the foreground indicated its type (monetary or erotic) and intensity (high or low, depending on the size of the pictogram). This led to a total of 12 different cues plus a control condition associated with no chance of winning (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). After a variable delay period (question mark representing a pseudorandom draw depending on probability), subjects were asked to perform a discrimination task, in which they had to respond correctly to a target within a maximum time of 1 s. The shape of the target was drawn at random on each trial and could be either a triangle (left button press) or a square (right button press). Success on this discrimination task (indicated by a magnified target) allowed the subjects to view the outcome of the pseudorandom draw, whereas erroneous or slow response (indicated by no change in target size) led to no reward. In rewarded trials, the reward was either an erotic image (with high or low erotic content) or the picture of a safe mentioning the amount of money won (high or low amount). After each reward outcome, subjects were asked to provide a hedonic rating by moving a cursor along a 1-to-9 continuous scale (1 for very little pleased; 9 for very highly pleased). In non-rewarded and control trials, the subjects were presented with “scrambled” pictures. A fixation cross was finally used as an intertrial interval of variable length.

Stimuli

Two categories (high and low intensity) of erotic pictures and monetary gains were used. Nudity being the main criteria driving the reward value of erotic stimuli, we separated them into a “low intensity” group displaying women in underwear or bathing suits and a “high intensity” group displaying naked women in an inviting posture. Each erotic picture was presented only once during the course of the task to avoid habituation. A similar element of surprise was introduced for the monetary rewards by randomly varying the amounts at stake: the low amounts were €1, €2, or €3 and the high amounts were €10, €11, or €12. The pictures displayed in non-rewarded and control trials were scrambled versions of the pictures used in rewarded trials and hence contained the same information in terms of chromaticity and luminance.

fMRI data acquisition

Imaging was conducted on a 1.5 T Siemens Sonata scanner, using an eight-channel head coil. The scanning session was divided into four runs. Each of them included four repetitions of each cue, with the exception of the control condition, repeated nine times. This yielded a total of 228 trials. Within each run, the order of the different conditions was pseudorandomized and optimized to improve signal deconvolution. The order of the runs was counterbalanced between subjects. Before scanning, all subjects were given oral instructions and familiarized with the cognitive task in a short training session.

Each of the four functional runs consisted of 296 volumes. Twenty-six interleaved slices parallel to the anterior commissure–posterior commissure line were acquired per volume (field of view, 220 mm; matrix, 64 \times 64; voxel size, 3.4 \times 3.4 \times 4 mm; gap, 0.4 mm), using a gradient-echo echoplanar imaging (EPI) T2*-weighted sequence (repetition time, 2500 ms; echo time, 60 ms; flip angle, 90°). To improve the local field homogeneity and hence minimize susceptibility artifacts in the orbitofrontal area, a manual shimming was performed within a rectangular region including the OFC and the basal ganglia. A high-resolution T1-weighted structural scan was subsequently acquired in each subject.

fMRI analysis

Preprocessing. Preprocessing of fMRI data was conducted using SPM2. The first four functional volumes of each run were removed, and the remaining images were corrected for slice-timing artifacts and spatially realigned to the first image of each time series. We then searched for residual artifacts in the time series with the *tsdiffana* utility (<http://imaging.mrc-cbu.cam.ac.uk/imaging/DataDiagnostics>) and modeled them with dummy regressors in our general linear model (two subjects

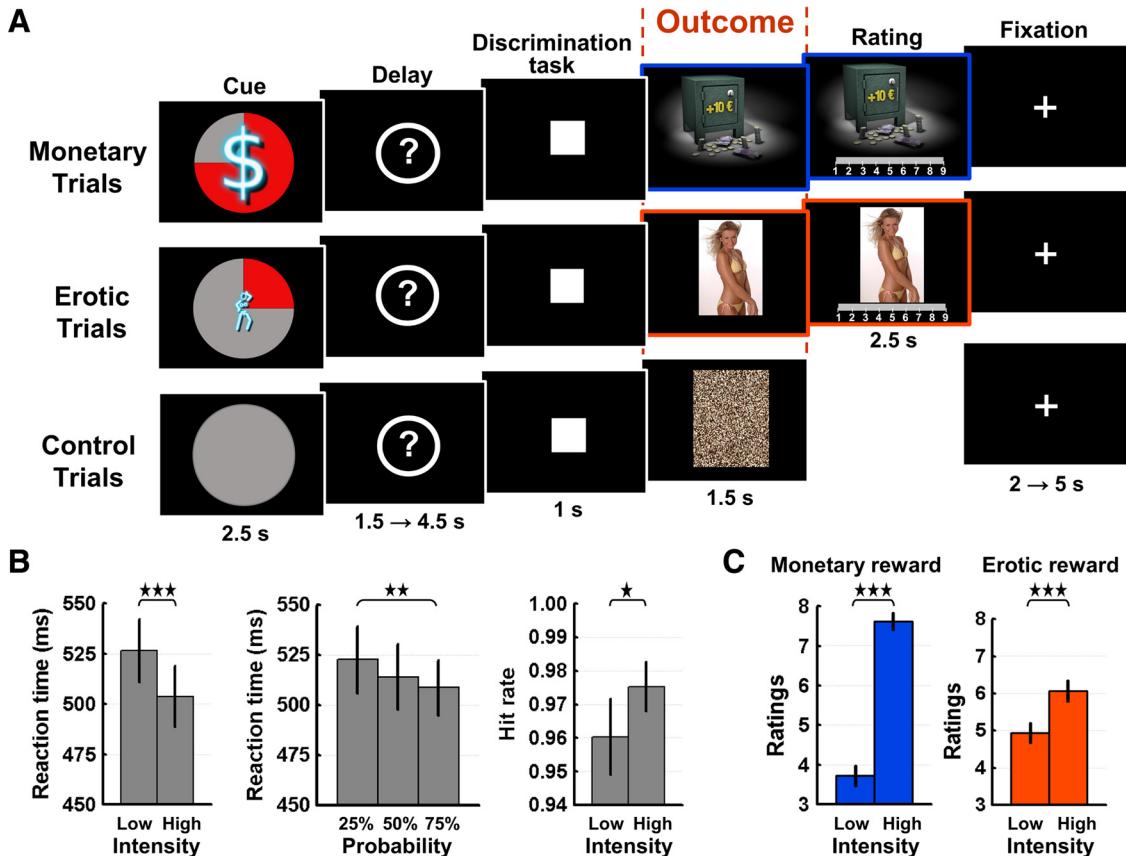


Figure 1. Paradigm and behavior. **A**, Sequence of events during a typical trial. Subjects first saw a cue informing them about the type, probability, and intensity of an upcoming reward (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Three cases are represented here: a 75% chance of receiving a high amount of money (top), a 25% chance of seeing a low erotic content picture (middle), and a sure chance of getting nothing (control trials; bottom). After a short delay and a target discrimination task, subjects saw the outcome, which was contingent on both the announced probability and their performance on the discrimination task. Reward outcomes consisted either in a monetary amount displayed on a safe (top) or an erotic picture (middle) and were followed by the rating of their subjective value on a continuous scale. Non-rewarded and control trials displayed a scrambled picture at outcome (bottom). **B**, Behavioral results on the discrimination task: mean reaction times according to reward intensity (left) and probability (middle) and mean hit rates according to reward intensity (right). **C**, Mean subjective ratings according to reward intensity, on a 1-to-9 scale. Error bars indicate SEM. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ by Tukey's HSD tests.

had three artifacts and one subject had one artifact in their time series). The functional images were then normalized to the Montreal Neurological Institute (MNI) stereotaxic space using the EPI template of SPM2 and spatially smoothed with a 10 mm full-width at half-maximum isotropic Gaussian kernel. Anatomical scans were normalized to the MNI space using the icbm152 template brain and averaged across subjects.

Identification of common and specific brain regions. The event-related statistical analysis was performed according to the general linear model as implemented in SPM2. Anticipation-related responses were modeled as boxcar functions time locked to the onset time of the cue with a duration of 2.5 s. The 2 rewards (monetary/erotic) \times 2 intensities (high/low) were modeled as four separate conditions. For each of them, a first-order parametric regressor modeled reward probability. The control condition was modeled in a separate regressor. Outcome-related responses were modeled as events time locked to the appearance of the reward (or scrambled picture). Four main conditions were defined: “monetary reward” (MR), “erotic reward” (ER), “no-monetary reward” (NoMR), and “no-erotic reward” (NoER). Two orthogonal covariates linearly modeling the expected probability and the ratings were added (in this order) to the MR and ER regressors. A last regressor modeled the appearance of a scrambled picture in the control condition (C). All regressors were subsequently convolved with the canonical hemodynamic response function and entered in a first-level analysis. A high-pass filter with a cutoff of 128 s was applied to the time series to remove low-frequency noise and baseline drifts. The resulting images of parameter estimates were then passed in a second-level group analysis in which between-subject variability was treated as a random effect.

Although the anticipatory period was explicitly modeled in our analysis, we only report results concerning the outcome phase because our focus was on the coding of experienced reward value. Following the forward inference approach (Henson, 2006) (see Results), brain regions responding specifically to monetary (or erotic) rewards resulted from the contrast MR > ER (or ER > MR), masked inclusively with MR > C (or ER > C) and exclusively with ER > C (or MR > C). The main contrasts MR > ER and ER > MR and the masks were thresholded independently using a whole-brain correction for multiple comparisons [$p < 0.05$ familywise error (FWE) and $p < 0.01$ false discovery rate (FDR), respectively], thereby ensuring the absence of any “selection bias” in the analysis (Kriegeskorte et al., 2009). Brain regions activated by both monetary and erotic rewards were identified in two steps. We first performed a conjunction analysis of the contrasts MR > C and ER > C based on the minimum statistic (Nichols et al., 2005) ($p < 0.05$ FWE whole-brain corrected). Because this conjunction may be sensitive to other types of computations (such as attention or image processing), we then masked it inclusively with the regions responding parametrically with both monetary and erotic hedonic ratings (each mask thresholded at $p < 0.05$ FDR whole-brain corrected).

Anatomical localization of functional clusters was performed based on a probabilistic atlas (Hammers et al., 2003).

Prediction error model. Positive prediction errors were defined at each trial t by $PE(t) = V(t) - B(t)$, where $V(t)$ is the outcome value and $B(t)$ is the expected value (Yacubian et al., 2006). Whereas monetary amounts could have been used to assess $V(t)$ for monetary rewards, erotic rewards could not be similarly quantified. Hence, to use an equivalent measure for both rewards, we used the hedonic rating to assess $V(t)$ on each trial.

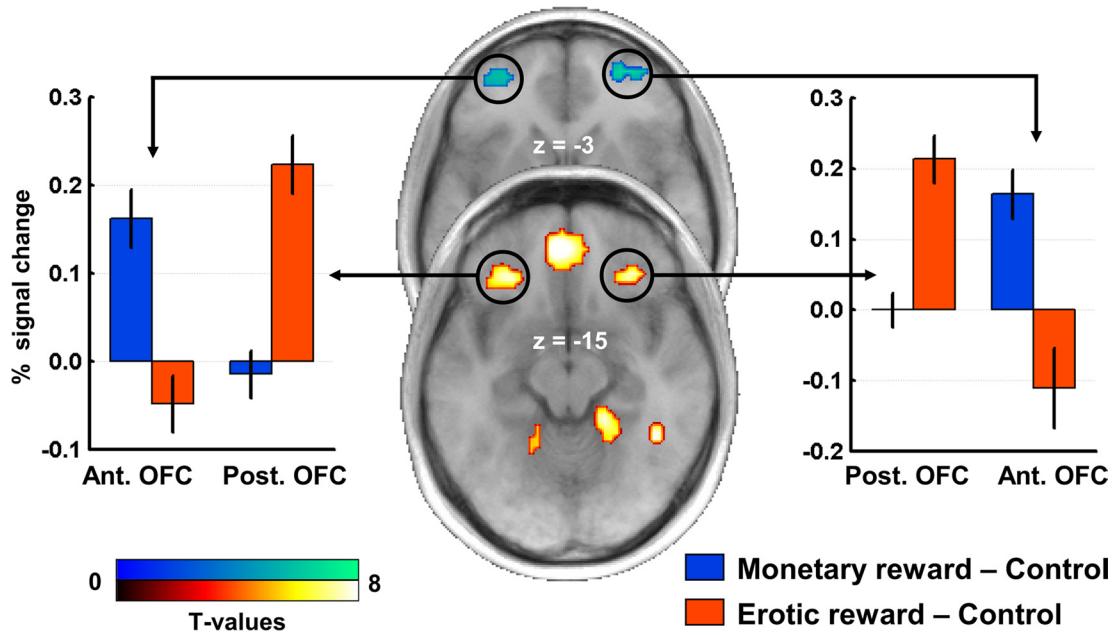


Figure 2. Functional postero-anterior dissociation in the orbitofrontal cortex depending on reward type. Brain regions responding specifically to monetary reward outcomes are displayed in blue–green, and those responding specifically to erotic reward outcomes are displayed in red–yellow. Plots of mean percent signal change, which are not independent of the whole-brain analysis, are shown only to illustrate the double dissociation between monetary/erotic rewards and anterior (Ant.)/posterior (Post.) OFC. Activations are overlaid on an average anatomical scan of all subjects ($p < 0.05$ FWE whole-brain corrected). Error bars indicate SEM.

$B(t)$ was defined as the product of reward probability $P(t)$ by expected intensity $E(t)$. $P(t)$ was simply the probability given explicitly in the cue (25, 50, or 75%). $E(t)$, presented as either “high” or “low” to the participants, was transformed into a numerical value by using the past ratings: for instance, $E(t)$ for a high monetary reward was estimated as the average of all the ratings given to high monetary rewards since the beginning of the task up to trial t .

Prediction error values were entered into two parametric regressors separately modeling monetary and erotic reward prediction errors (PE_{MR} and PE_{ER}). Note, however, that prediction error was relatively correlated with outcome value, i.e., with the hedonic ratings (mean $r = 0.72$ for monetary rewards and mean $r = 0.75$ for erotic rewards). This correlation is inherent to the nature of these signals and is a classical shortcoming of fMRI studies on reward processing (Hare et al., 2008). As a consequence, if prediction errors and hedonic ratings were entered in the same general linear model, they would both end up with a rather low explanatory power, because only the orthogonal component of each regressor would be allowed to compete for variance (Hunt, 2008) (such a model was estimated and produced poor results in expected brain regions such as the ventral striatum and the OFC). For this reason, we built a separate general linear model, in which reward outcomes were modulated only by prediction errors (instead of probability and ratings). We should emphasize, however, that this procedure makes it difficult to distinguish the contribution of prediction error and outcome value computations at the brain level.

The resulting T -maps showing positive correlations between the blood oxygenation level-dependent (BOLD) signal and monetary or erotic prediction errors were subsequently entered in a conjunction analysis (Nichols et al., 2005) thresholded at $p < 0.001$ uncorrected for multiple comparisons at the voxel level.

Region of interest analyses. Region of interest (ROI) analyses were conducted with the extension of SPM MarsBaR (<http://marsbar.sourceforge.net/>) within ROIs defined functionally from the whole-brain analyses. Each ROI was created by taking the intersection of the functional cluster of interest and a 10-mm-radius sphere centered on the highest peak voxel of the cluster (to isolate distinct brain areas pertaining to the same cluster). In keeping with the approach of MarsBaR, percent signal change for a given condition in a given ROI was calculated as the effect size of that

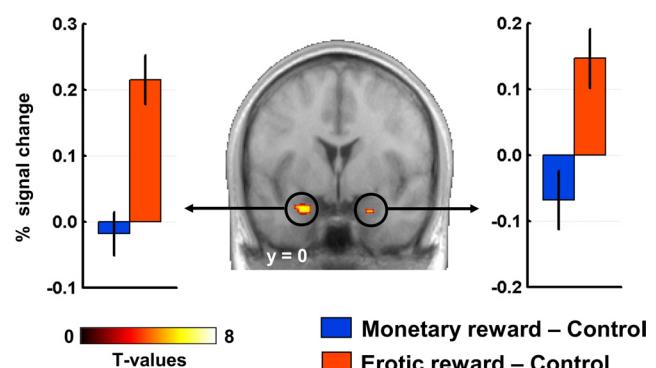


Figure 3. Specific response of amygdala to erotic rewards. Activations are overlaid on an average anatomical scan of all subjects ($p < 0.05$ FWE whole-brain corrected). Left and right plots of mean percent signal change, which are not independent of the whole-brain analysis, are shown only to illustrate the specificity of amygdalar response. Error bars indicate SEM.

condition (β value) divided by the mean activity of that ROI and multiplied by 100.

Results

Behavior

Hit rates and reaction times (RTs), obtained at the time of the discrimination task, as well as hedonic ratings obtained at the time of outcome, were analyzed in separate three-way ANOVAs including reward type, probability, and intensity as within-subject factors. The analysis on hit rates and RT was performed on 17 subjects only, because data were accidentally lost for one subject. The mean hit rate across subjects on the discrimination task was 96%.

There was no significant main effect of reward type on hit rates ($p = 0.38$) and RT ($p = 0.20$), suggesting that monetary gains and erotic pictures had comparable incentive values.

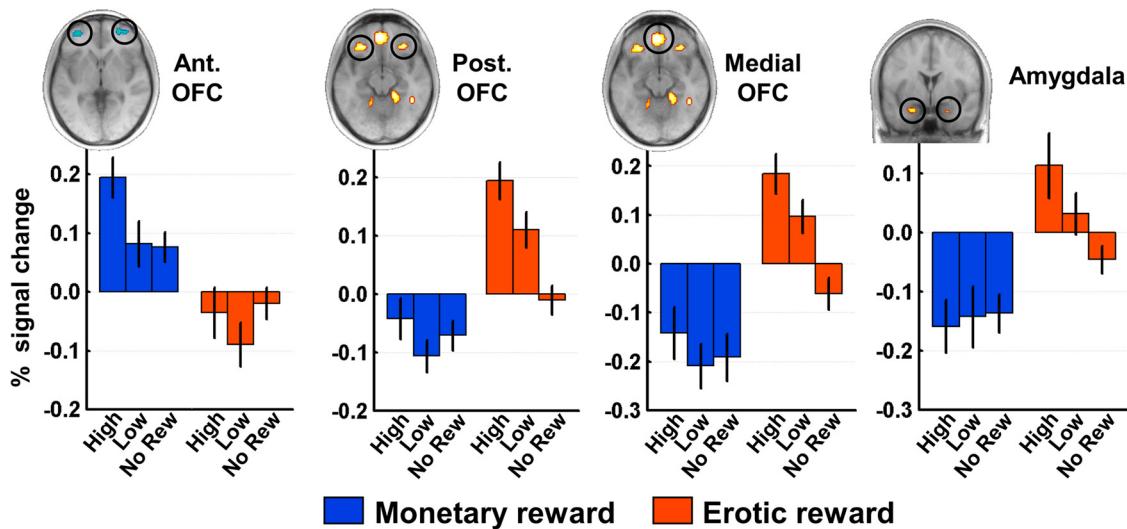


Figure 4. Response pattern of the reward-specific brain regions as a function of reward intensity. Percent signal change is plotted in the circled ROIs for monetary and erotic rewards according to the following conditions: high intensity, low intensity, and no reward. In each region, brain activity increases with reward intensity only for the reward for which it is specific. Error bars indicate SEM. The signal is averaged across the right and left hemispheres in each brain region (similar patterns of activity were observed in each hemisphere). Ant., Anterior; Post., posterior.

Subjects were faster ($F_{(1,16)} = 34.2, p < 0.001$) and more accurate ($F_{(1,16)} = 7.7, p < 0.05$) for high intensity incentives and were also faster for more likely rewards ($F_{(2,32)} = 5.3, p < 0.05$) (Fig. 1B). These results reflect increased motivation for higher reward intensity and more certain rewards. They also confirm that subjects were engaged in the task and effectively encoded the cue information. Importantly, these effects were similar for monetary and erotic rewards, as shown by the absence of significant interaction between intensity and reward type (hit rate, $p = 0.67$; RT, $p = 0.20$) (supplemental Fig. 2, available at www.jneurosci.org as supplemental material) and between probability and reward type (RT, $p = 0.11$).

No significant effect of reward type was observed on the hedonic ratings ($p = 0.40$), suggesting that monetary and erotic rewards had similar subjective values. Conversely, we found a robust main effect of intensity on the ratings ($F_{(1,17)} = 150.8, p < 0.001$), which remained significant for each type of reward taken separately [Tukey's honestly significant difference (HSD) tests: monetary rewards, $T_{(17)} = 20.4, p < 0.001$; erotic rewards, $T_{(17)} = 4.4, p < 0.001$] (Fig. 1C). This shows that, for both rewards, the two intensity categories chosen *a priori* (high vs low) were effectively perceived by the subjects. The ratings also showed an interaction between reward type and intensity ($F_{(1,17)} = 111.5, p < 0.001$), simply because the subjects used a smaller portion of the scale to rate erotic pictures. Finally, the ratings were not influenced by reward probability ($F_{(2,34)} = 1.4, p = 0.26$), confirming that they reflected a purely hedonic evaluation.

Neuroimaging data

Reward-specific brain regions

Brain regions specific for each type of reward were identified based on the forward inference approach proposed by Henson (2006) to demonstrate qualitative differences in brain imaging data. Specifically, money-specific regions were defined as those stemming from the comparison MR > ER, masked inclusively with MR > C and exclusively with ER > C, and conversely erotic-specific regions were defined as those responding in the comparison ER > MR, masked inclusively with ER > C and exclusively with MR > C. This procedure ensures that the resulting brain areas

meet two criteria: (1) they are more activated by one reward compared with the other (“dissociation” criterion); (2) they respond to either monetary or erotic rewards, but not to both, compared with a common control condition (“association” criterion).

As hypothesized, monetary rewards specifically recruited the anterior lateral OFC (MNI [x y z] $[-30, 51, 0]$, $T = 5.92$; $[30, 54, -3]$, $T = 6.80$), spanning the anterior orbital gyrus, the lateral orbital gyrus, and the ventral part of the middle frontal gyrus (Fig. 2). In contrast, erotic rewards elicited activity specifically in the posterior part of the lateral OFC ($[-30, 33, -15]$, $T = 7.56$; $[30, 33, -15]$, $T = 7.54$), straddling the posterior and lateral orbital gyri (Fig. 2). These results demonstrate a double dissociation between monetary/erotic rewards and the anterior/posterior OFC, which is further illustrated in the bar graphs of Figure 2, representing the MR > C and ER > C differences in percent signal change extracted from these regions. Among erotic-specific areas, a large cluster was also present in the medial OFC ($[-6, 45, -15]$, $T = 8.90$), encompassing the medial orbital gyrus, the straight gyrus, and the most ventral part of the superior frontal gyrus (Fig. 2). Subcortically, the only structure specifically activated by erotic pictures was the bilateral amygdala ($[-21, -6, -27]$, $T = 6.94$; $[24, 0, -27]$, $T = 5.45$) (Fig. 3). Other money-specific and erotic-specific foci are reported in supplemental Tables 1 and 2 (available at www.jneurosci.org as supplemental material), respectively.

This segregated representation of reward types in the OFC was not merely attributable to visual or hedonic differences between monetary and erotic outcomes, because an identical dissociation emerged when we repeated the same analysis using no-reward outcomes instead of reward outcomes. That is, when comparing the NoMR and NoER conditions, which only differ with respect to the type of reward being expected (while comparing visually identical scrambled pictures), the anterior OFC specifically responded to no-monetary outcomes ($[-30, 51, 0]$, $T = 6.16$; $[33, 51, 0]$, $T = 6.20$), whereas the posterior OFC specifically responded to no-erotic outcomes ($[-21, 33, -12]$, $T = 6.39$) (supplemental Fig. 3, supplemental Tables 1, 2, available at www.jneurosci.org as supplemental material). This result excludes a mere

perceptual account of the functional dissociation observed in the OFC and supports the idea that monetary and erotic rewards are encoded in distinct OFC regions.

Moreover, ROI analyses on reward intensity coding brought additional evidence in support of a segregated representation of monetary and erotic outcomes. For both rewards, we extracted the percent signal change for the high reward, low reward, and no-reward conditions (Fig. 4). The results show that, whereas activity in the anterior OFC increased with monetary reward intensity, such a monotonic variation was not present for increasing levels of erotic reward intensity. Conversely, in the posterior OFC, medial OFC, and amygdala, activity was found to increase monotonically with erotic reward intensity but not with monetary reward intensity. Together, these findings indicate that reward-specific brain regions only reflected intensity for the reward type they specifically encoded.

Finally, to further confirm the specificity of these regions with respect to the subjective experience of monetary and erotic outcomes, we performed a whole-brain analysis comparing the coding of their hedonic value. For each reinforcer, reward and no-reward outcomes were pooled together, while the hedonic value was modeled in a parametric regressor with the corresponding continuous rating (from 1 to 9) or a 0, respectively. As expected, the contrast of these parametric regressors between the two rewards revealed the same brain regions as the previous categorical analysis. In particular, BOLD activity in the right anterior lateral OFC was found to scale best with monetary hedonic value, whereas activity in the posterior lateral OFC, medial OFC, and amygdala was found to scale best with erotic hedonic value (supplemental Fig. 4, available at www.jneurosci.org as supplemental material).

Common reward brain regions

To identify the brain regions commonly activated by monetary and erotic outcomes, we first compared each reward with the control condition and performed a conjunction of these two comparisons (supplemental Fig. 5, available at www.jneurosci.org as supplemental material). Because this analysis remains sensitive to non-reward-related computations such as attentional or image processing effects, it was masked by the brain regions responding parametrically with the subjective value of both monetary and erotic rewards (i.e., hedonic ratings) (Fig. 5). This procedure revealed significant bilateral activations in a set of brain regions classically involved in reward processing: the ventral striatum ($[−12, 9, −9], T = 6.38; [9, 6, −9], T = 6.20$), the midbrain ($[−3, −24, −24], T = 7.30$), the anterior cingulate cortex (ACC) ($[−6, 27, 39], T = 8.15; [9, 18,$

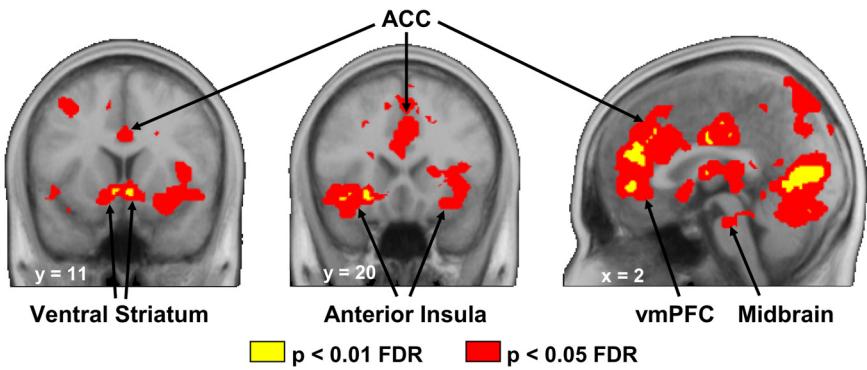


Figure 5. Brain regions reflecting hedonic ratings regardless of reward type. Activations show the brain areas in which activity positively correlates with both monetary and erotic ratings (intersection of T -maps thresholded at $p < 0.01$ FDR whole-brain corrected shown in yellow or thresholded at $p < 0.05$ FDR whole-brain corrected shown in red). The regions displayed in red were used as an inclusive mask in the analysis of Figure 6. Activations are overlaid on an average anatomical scan of all subjects.

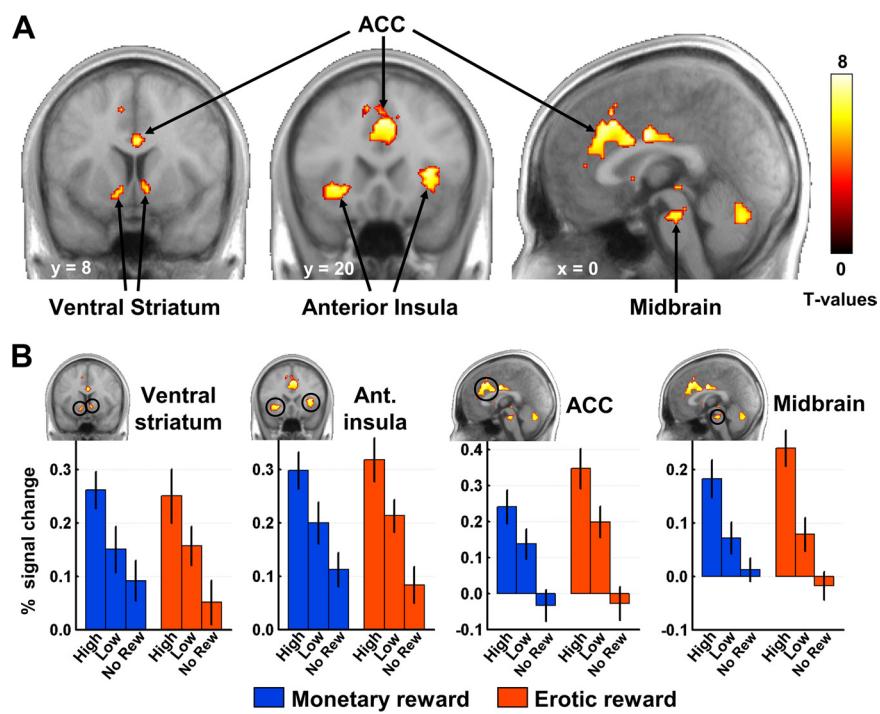


Figure 6. Common reward brain regions. **A**, T -map showing the brain regions encoding experienced reward value for both monetary and erotic reward outcomes. Activations are overlaid on an average anatomical scan of all subjects ($p < 0.05$ FWE whole-brain corrected). **B**, Percent signal change is plotted in the circled ROIs for monetary and erotic rewards according to the following conditions: high intensity, low intensity, and no reward. Note that these plots are not independent of the whole-brain analysis and are only shown as an illustration for easier visual comparison with Figure 4. Error bars indicate SEM. The signal is averaged across the right and left hemispheres in each brain region (similar patterns of activity were observed in each hemisphere). Ant., Anterior.

39], $T = 8.42$), and the anterior insula ($[−27, 21, −6], T = 7.48; [33, 24, 3], T = 8.14$) (Fig. 6A). Other foci are reported in supplemental Table 3 (available at www.jneurosci.org as supplemental material).

The relative coding of reward intensity in these regions was further illustrated using the same ROI approach as with reward-specific regions (Fig. 6B). Note, however, that this representation is not independent of the previous whole-brain analysis, because reward intensity is highly correlated with the hedonic ratings that have served in the identification of the “common network.” Consequently, the resulting bar graphs are purely illustrative of the conclusion drawn from the T -map, which is that activity in the

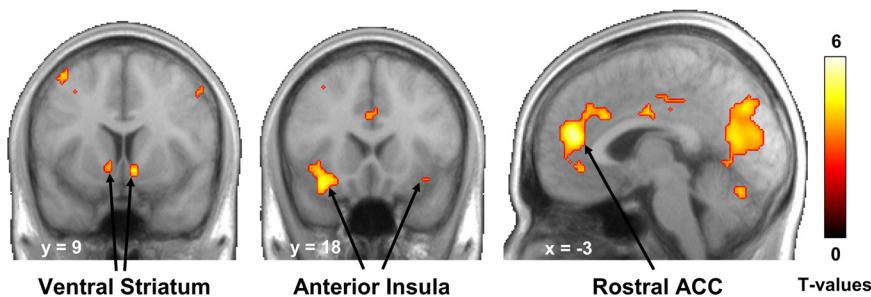


Figure 7. Brain regions reflecting prediction errors regardless of reward type. Activations result from a conjunction analysis showing the brain regions in which activity positively correlates with both monetary and erotic prediction errors ($p < 0.001$ uncorrected). Activations are overlaid on an average anatomical scan of all subjects.

ventral striatum, midbrain, ACC, and anterior insula reflects experienced reward value regardless of reward type.

Coding of reward prediction errors

Reward-related brain regions are thought to mediate prediction error signals, coding the difference between expected outcomes and those effectively delivered (McClure et al., 2003; O'Doherty et al., 2003b; Schultz, 2006). These prediction error signals occur not only in conditioning procedures and can also be computed in non-learning situations, such as in the present study (Dreher et al., 2006; Yacubian et al., 2006). To determine whether such signals are supported by similar brain networks for monetary and erotic rewards, the fMRI data were fitted with parametric regressors modeling positive prediction errors related to reward outcomes. Our results revealed that monetary and erotic reward prediction errors were processed in a similar set of brain regions, essentially overlapping with the previously identified common network. Specifically, a conjunction analysis showed that activity in the ventral striatum ($[-9, 9, -6], T = 3.98$; $[6, 9, -9], T = 4.57$), anterior insula ($[-33, 18, -18], T = 4.80$; $[39, 21, -15], T = 3.37$), and rostral ACC ($[-3, 42, 15], T = 5.57$) correlated positively with prediction errors regardless of reward type (Fig. 7) (supplemental Table 4, available at www.jneurosci.org as supplemental material). Moreover, the comparison of monetary and erotic reward prediction errors revealed almost no activity in reward-related areas (at $p < 0.001$ uncorrected, only a tiny cluster appeared in the left pallidum for the contrast of erotic minus monetary prediction errors) and especially not in the brain regions labeled as “reward specific” (even at a liberal threshold of $p < 0.05$).

Discussion

As predicted, the OFC was found to be functionally organized along a postero-anterior axis with respect to reward type, with the anterior part responding exclusively to money and the posterior part responding exclusively to erotic stimuli. Additional erotic-specific activations were also found in the bilateral amygdala and medial OFC. Importantly, brain activity in these reward-specific regions only scaled with the hedonic value of the reward they specifically encoded. In parallel, our results support the idea of a core reward system processing experienced rewards regardless of their nature. In this common network, including the ventral striatum, ACC, anterior insula, and midbrain, functional activity correlated with hedonic value and prediction error for both monetary gains and erotic pictures. Together, our results reveal the existence of both reward-specific and nonspecific brain networks, challenging the view of a unique reward system for all reinforcers.

Postero-anterior dissociation in the orbitofrontal cortex

The segregated responses to monetary and erotic outcomes along a postero-anterior axis in the OFC suggest a functional division of experienced reward value representation according to an abstractness gradient. Paralleling this result in the domain of cognitive control, recent theories on the functional divisions of the human lateral prefrontal cortex proposed that it is organized hierarchically, whereby cognitive control involving temporally proximate and concrete action representations is supported by posterior lateral prefrontal regions, and cognitive control involving temporally extended and abstract representations is supported by more anterior lateral prefrontal regions such as the frontopolar cortex (Koechlin and Summerfield, 2007; Badre, 2008; Dreher et al., 2008a). Our study shows that a similar unifying principle of caudo-rostral hierarchical organization can be applied to the OFC. Notably, patients with lesions in the anterior OFC have been reported to be specifically impaired in making decisions entailing abstract, i.e., distant, consequences, and not in making decisions leading to concrete, i.e., immediate, consequences, further supporting a postero-anterior trend in the representation of abstractness in the OFC (Bechara and Damasio, 2005).

Although an anatomical gradient in the postero-anterior axis of the OFC had been suggested based on cytoarchitectonic data (Ongür and Price, 2000; Wise, 2008), its functional relevance for reward processing had never been tested empirically. Our data bring strong empirical support to this hypothesis. Although a generalization to all primary and secondary rewards cannot be ascertained in a single fMRI study, it is consistent with the pattern of activation observed in the recent literature for other rewards (Kringelbach and Rolls, 2004). In particular, the medial and posterior lateral OFC, responding to erotic pictures in the current and previous studies (Ponseti et al., 2006), were shown to respond to other primary rewards, such as attractive faces (O'Doherty et al., 2003a), pleasant odors (Gottfried et al., 2006), and pleasant taste (Small et al., 2001). Conversely, other studies manipulating monetary (Reuter et al., 2005; Vollm et al., 2007) or social (Izuma et al., 2008) rewards have reported a similar anterior OFC region as the one we found.

One intrinsic property of erotic pictures is that they are the reward, whereas monetary rewards delivered in the scanner are a representation of what the participant will receive at the end of the experiment. Thus, it could be argued that the OFC dissociation relates to the immediate rewarding effect of erotic pictures compared with the delayed rewarding effect of monetary gains. This is unlikely to be the case, because (1) our pattern of OFC activation was not observed for immediate versus delayed rewards in intertemporal choice studies (Kable and Glimcher, 2007; Prevost et al., 2010) and (2) the same OFC functional dissociation emerged when monetary and erotic rewards were expected but not effectively delivered. This finding also suggests that the segregated representation of reward types in the OFC was not merely attributable to visual or intrinsic differences between monetary and erotic outcomes, such as saliency or arousal.

Electrophysiological recordings in monkeys indicate that OFC neurons encode the economic value assigned to different rewarding juices when choosing between them (Padoa-Schioppa and Assad, 2008). Some OFC neurons were also found to encode taste responses reflecting the identity of a chosen juice. However, no neuronal recording study has yet investigated whether primary and secondary rewards (e.g., juice vs social dominance) are coded in distinct OFC subregions. Our finding of a clear OFC dissociation in humans supports a hierarchical organization along a continuum from the posterior to the anterior part of the OFC. Whether and how these distinct representations of value in the OFC are preserved across species remains an important open question. Our results also suggest that learning through secondary reinforcement may depend more on anterior OFC regions, whereas primary reinforcement may depend more on the posterior OFC. This hypothesis, which remains to be tested, may shed light on studies across a range of animal species indicating effects of OFC lesions on behavior maintained or acquired through secondary reinforcement (Murray et al., 2007).

In addition, our findings clarify how value signals in the OFC are integrated with those from other brain structures. Together with the posterior and medial OFC, we found that the amygdala responded exclusively to erotic pictures. Previous neuroimaging studies also reported that erotic pictures evoke amygdala response (Redouté et al., 2000; Karama et al., 2002; Hamann et al., 2004), whereas money, as a secondary reinforcer, often failed to do so in studies using experimental designs similar to the present paradigm (Knutson et al., 2001). This is consistent with the underlying anatomy, showing that the amygdala is more connected with the posterior and medial OFC than with the anterior OFC (Carmichael and Price, 1995). Moreover, the parametric modulation of the amygdala with erotic hedonic value is in accordance with its general role in emotional arousal for both appetitive and aversive stimuli. However, amygdala response to erotic stimuli is not solely determined by arousal, as suggested by a previous study reporting higher amygdala response in men than in women viewing sexual stimuli, despite similar arousal ratings in both groups (Hamann et al., 2004).

Brain regions common to monetary and erotic rewards

An important strength of our experimental design, compared with previous reward studies, is that it made it possible to directly test whether monetary and erotic reward outcomes are truly encoded within the same brain regions. The enhanced response to increasing hedonic value observed in the common network regardless of reward type suggests that this network processes experienced reward value in a general manner. This is consistent with its reported implication (in separate studies) in the hedonic processing of rewards ranging from primary reinforcers, such as sexual stimuli (Redouté et al., 2000), attractive faces (Bray and O'Doherty, 2007; Smith et al., 2010), and pleasant taste (Small et al., 2001), to secondary rewards, such as money or social approval (Izuma et al., 2008). Our common network activity is also compatible with an interpretation in terms of general arousal or saliency (Zink et al., 2004).

To efficiently compare the goal values of different rewards during decision making, it has been proposed that the brain may convert them into a common neural currency (Montague and Berns, 2002). A wealth of electrophysiological and fMRI studies has since confirmed this hypothesis, emphasizing in particular the role of the ventral striatum and vmPFC (Kable and Glimcher,

2007; Knutson et al., 2007; Chib et al., 2009). The present results suggest that a similar computation might be performed in the same brain network at the time of reward consumption. Indeed, along with the ventral striatum, activity in the vmPFC also reflected the hedonic experience of the participants regardless of reward type (Fig. 5). Such a mechanism encoding heterogeneous outcome values on a common scale might be helpful for efficient comparison of these values during subsequent value-based decision making.

Finally, we found prediction-error-related activity in the ventral striatum, anterior insula, and ACC for both monetary and erotic stimuli. This finding demonstrates that prediction errors are computed in a network independent of reward type and generalizes this concept to the domain of erotic stimuli, paralleling previous work performed with monetary gains (Yacubian et al., 2006), pleasant taste (O'Doherty et al., 2003b), and attractive faces (Bray and O'Doherty, 2007). Thus, prediction errors may be a primitive neural signal computed in midbrain dopaminergic neurons regardless of reward type and primarily delivered to the common reward network that may be responsible for making predictions. However, it is important to emphasize that, because of the inherent correlation existing between prediction error and hedonic value, it is difficult to disentangle these computations at the brain level (Behrens et al., 2008; Hare et al., 2008). Consequently, although the response pattern observed in our common network is consistent with both interpretations, we cannot ascribe with certainty one role or the other to these regions. Concerning the ventral striatum, however, a recent study aiming to dissociate reward value from prediction error found that activity in this region best correlated with the latter (Hare et al., 2008). Although this study focused on the computation of goal values in the context of decision making, and therefore cannot be directly compared with the present one, it brings evidence favoring the prediction error hypothesis in the ventral striatum. Note that midbrain activity was observed in our common network, but not in the prediction error analysis, whereas single-neuron recordings classically report prediction error signals in midbrain dopaminergic structures (Schultz, 2006). Although further work is needed to gain a better understanding of the relationship between dopaminergic neuron firing and the BOLD signal observed in reward paradigms, recent findings combining fMRI with FDOPA positron emission tomography measures of midbrain dopamine synthesis (Dreher et al., 2008b) and analyses of variations in genes involved in dopamine transmission established a link between higher prefronto-striatal BOLD signal and dopamine synaptic availability during reward processing (Dreher et al., 2009).

Conclusion

Our results provide plausible functional mechanisms explaining the existence of two separate reward networks in the brain. The nature of their interactions remains to be determined, but one possibility is that outcome value signals computed in the reward-specific OFC regions would be sent to the common network for additional integration and comparison processes. From an evolutionary perspective, the distinct cytoarchitectonic properties of the anterior and posterior parts of the OFC suggest that the ability to process primary rewards may occur phylogenetically and ontogenetically earlier than the ability to process secondary rewards, which represent more evolved adaptive behavior. Our findings also have important clinical implications for a range of neuropsychopathological disorders characterized by major deficits in mo-

tivation and behavioral control, such as pathological gambling or hypersexuality. The dissociable representation of various rewards along a postero-anterior axis in the OFC may shed light on this important question.

References

- Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage* 31:790–795.
- Badre D (2008) Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn Sci* 12:193–200.
- Bechara A, Damasio A (2005) The somatic marker hypothesis: A neural theory of economic decision. *Games Econ Behav* 52:336–372.
- Beck AT, Beck RW (1972) Screening depressed patients in family practice. A rapid technic. *Postgrad Med* 52:81–85.
- Behrens TE, Hunt LT, Woolrich MW, Rushworth MF (2008) Associative learning of social value. *Nature* 456:245–249.
- Bray S, O'Doherty J (2007) Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *J Neurophysiol* 97:3036–3045.
- Caldú X, Dreher JC (2007) Hormonal and genetic influences on processing reward and social information. *Ann NY Acad Sci* 1118:43–73.
- Carmichael ST, Price JL (1995) Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* 363:615–641.
- Chib VS, Rangel A, Shimojo S, O'Doherty JP (2009) Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci* 29:12315–12320.
- Deaner RO, Khera AV, Platt ML (2005) Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr Biol* 15:543–548.
- Dreher JC (2009) Decomposing brain signals involved in value-based decision making. In: *Handbook of reward and decision making* (Dreher JC, Tremblay L, eds), pp 137–164. New York: Academic/Elsevier.
- Dreher JC, Kohn P, Berman KF (2006) Neural coding of distinct statistical properties of reward information in humans. *Cereb Cortex* 16:561–573.
- Dreher JC, Schmidt PJ, Kohn P, Furman D, Rubinow D, Berman KF (2007) Menstrual cycle phase modulates reward-related neural function in women. *Proc Natl Acad Sci U S A* 104:2465–2470.
- Dreher JC, Koechlin E, Tierney M, Grafman J (2008a) Damage to the fronto-polar cortex is associated with impaired multitasking. *PLoS One* 3:e3227.
- Dreher JC, Meyer-Lindenberg A, Kohn P, Berman KF (2008b) Age-related changes in midbrain dopaminergic regulation of the human reward system. *Proc Natl Acad Sci U S A* 105:15106–15111.
- Dreher JC, Kohn P, Kolachana B, Weinberger DR, Berner KF (2009) Variation in dopamine genes influences responsiveness of the human reward system. *Proc Natl Acad Sci U S A* 106:617–622.
- FitzGerald TH, Seymour B, Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J Neurosci* 29:8388–8395.
- Gomes CM, Boesch C (2009) Wild chimpanzees exchange meat for sex on a long-term basis. *PLoS One* 4:e5116.
- Gottfried JA, Small DM, Zald DH (2006) The chemical senses. The orbitofrontal cortex, pp 125–171. Oxford: Oxford UP.
- Hamann S, Herman RA, Nolan CL, Wallen K (2004) Men and women differ in amygdala response to visual sexual stimuli. *Nat Neurosci* 7:411–416.
- Hammers A, Allom R, Koepf MJ, Free SL, Myers R, Lemieux L, Mitchell TN, Brooks DJ, Duncan JS (2003) Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* 19:224–247.
- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 28:5623–5630.
- Hayden BY, Parikh PC, Deane RO, Platt ML (2007) Economic principles motivating social attention in humans. *Proc Biol Sci* 274:1751–1756.
- Henson R (2006) Forward inference using functional neuroimaging: dissociations versus associations. *Trends Cogn Sci* 10:64–69.
- Hoon E, Chambliss D (1998) Sexual arousability inventory and sexual arousability inventory—expanded. In: *Handbook of Sexuality-Related Measures*, pp 71–74. Washington, DC: American Psychiatric Association.
- Hunt LT (2008) Distinctive roles for the ventral striatum and ventral prefrontal cortex during decision-making. *J Neurosci* 28:8658–8659.
- Izuma K, Saito DN, Sadato N (2008) Processing of social and monetary rewards in the human striatum. *Neuron* 58:284–294.
- Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10:1625–1633.
- Karama S, Lecours AR, Leroux JM, Bourguin P, Beaudoin G, Joubert S, Beauregard M (2002) Areas of brain activation in males and females during viewing of erotic film excerpts. *Hum Brain Mapp* 16:1–13.
- Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001) Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport* 12:3683–3687.
- Knutson B, Taylor J, Kaufman M, Peterson R, Glover G (2005) Distributed neural representation of expected value. *J Neurosci* 25:4806–4812.
- Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G (2007) Neural predictors of purchases. *Neuron* 53:147–156.
- Koechlin E, Summerfield C (2007) An information theoretical approach to prefrontal executive function. *Trends Cogn Sci* 11:229–235.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540.
- Kringelbach ML, Rolls ET (2004) The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog Neurobiol* 72:341–372.
- McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- McClure SM, Ericson KM, Laibson DI, Loewenstein G, Cohen JD (2007) Time discounting for primary rewards. *J Neurosci* 27:5796–5804.
- Montague PR, Berns GS (2002) Neural economics and the biological substrates of valuation. *Neuron* 36:265–284.
- Murray EA, O'Doherty JP, Schoenbaum G (2007) What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *J Neurosci* 27:8166–8169.
- Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005) Valid conjunction inference with the minimum statistic. *Neuroimage* 25:653–660.
- O'Doherty J, Winston J, Critchley H, Perrett D, Burt DM, Dolan RJ (2003a) Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia* 41:147–155.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003b) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- Onogür D, Price JL (2000) The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* 10:206–219.
- Padoa-Schioppa C, Assad JA (2008) The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat Neurosci* 11:95–102.
- Peters J, Büchel C (2010) Neural representations of subjective reward value. *Behav Brain Res* 213:135–141.
- Plassmann H, O'Doherty J, Rangel A (2007) Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* 27:9984–9988.
- Ponseti J, Bosinski HA, Wolff S, Peller M, Jansen O, Mehdorn HM, Büchel C, Siebner HR (2006) A functional endophenotype for sexual orientation in humans. *Neuroimage* 33:825–833.
- Prevost C, Pessiglione M, Météreau E, Clery-Melin ML, Dreher JC (2010) Distinct valuation subsystems in the human brain for effort and delay. *J Neurosci*, in press.
- Rangel A, Camerer C, Montague PR (2008) A framework for studying the neurobiology of value-based decision making. *Nat Rev Neurosci* 9:545–556.
- Redouté J, Stoléru S, Grégoire MC, Costes N, Cinotti L, Lavenne F, Le Bars D, Forest MG, Pujol JF (2000) Brain processing of visual sexual stimuli in human males. *Hum Brain Mapp* 11:162–177.
- Reuter J, Raedler T, Rose M, Hand I, Gläscher J, Büchel C (2005) Pathological gambling is linked to reduced activation of the mesolimbic reward system. *Nat Neurosci* 8:147–148.
- Reynolds CF 3rd, Frank E, Thase ME, Houck PR, Jennings JR, Howell JR, Lilienfeld SO, Kupfer DJ (1988) Assessment of sexual function in depressed, impotent, and healthy men: factor analysis of a brief sexual function questionnaire for men. *Psychiatry Res* 24:231–250.
- Rhodes G (2006) The evolutionary psychology of facial beauty. *Annu Rev Psychol* 57:199–226.

- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115.
- Small DM, Zatorre RJ, Dagher A, Evans AC, Jones-Gotman M (2001) Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain* 124:1720–1733.
- Smith DV, Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA (2010) Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J Neurosci* 30:2490–2495.
- Sugrue LP, Corrado GS, Newsome WT (2005) Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat Rev Neurosci* 6:363–375.
- Thornhill R, Gangestad SW (1999) Facial attractiveness. *Trends Cogn Sci* 3:452–460.
- Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398:704–708.
- Völlm B, Richardson P, McKie S, Elliott R, Dolan M, Deakin B (2007) Neuronal correlates of reward and loss in Cluster B personality disorders: a functional magnetic resonance imaging study. *Psychiatry Res* 156:151–167.
- Wise SP (2008) Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci* 31:599–608.
- Yacubian J, Gläscher J, Schroeder K, Sommer T, Braus DF, Büchel C (2006) Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J Neurosci* 26:9530–9537.
- Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS (2004) Human striatal responses to monetary reward depend on saliency. *Neuron* 42:509–517.

DECISION MAKING



So how do you really make up your mind? Are you always on top of your decisions? How much unconscious influence is there on your actual choice? What happens in the case of doubt?

After reading the previous sections, you will now understand that decision making is the result of a number of previous complex steps. What made you attend to some things rather than others? How did you respond emotionally? Were any memories triggered?

When we employ neuroscience methods, we get a much better understanding of the actual processes underlying a choice. We see that branding works through a series of unconscious and conscious operations; we find that valuation of choice options occurs in a fraction of a second; and we see that the reasons people report for making a choice do not always correspond to what actually happened.

Predicting Purchase Decision: The Role of Hemispheric Asymmetry Over the Frontal Cortex

Niklas Ravaja

Aalto University, Helsinki Institute for Information
Technology (HIIT), and University of Helsinki

Outi Somervuori and Mikko Salminen

Aalto University

This article examines how approach motivation as indexed by electroencephalographic (EEG) asymmetry over the prefrontal cortex predicts purchase decision when brand and price are varied. In a within-subjects design, the participants were presented purchase decision trials with 14 different grocery products (seven private label and seven national brand products) whose prices were increased and decreased while their EEG activity was recorded. The results showed that relatively greater left frontal activation (i.e., higher approach motivation) during the predecision period predicted an affirmative purchase decision. The relationship of frontal EEG asymmetry with purchase decision was stronger for national brand products compared with private label products and when the price of a product was below a normal price (i.e., implicit reference price) compared with when it was above a normal price. Higher perceived need for a product and higher perceived product quality were associated with greater relative left frontal activation.

Keywords: purchase decision, price, brand, electroencephalography, neurophysiology

According to the neoclassical view of a rational *Homo Economicus*, humans make choices based on rational Bayesian maximization of expected utility, as if they were equipped with unlimited knowledge, time, and information-processing power (Naqvi, Shiv, & Bechara, 2006; Oullier, Kirman, & Kelso, 2008). This view has been challenged, for example, by the Prospect Theory stating that subjective utility is dependent on a reference point and that people tend to strongly prefer avoiding losses over acquiring gains (Kahneman & Tversky, 1979). A mounting body of evidence shows that emotional processes play a crucial role in economic decision making (e.g., Bernheim & Rangel, 2004; Kahneman, Ritov, & Schkade, 1999; Loewenstein & Lerner, 2003; Shiv & Fe-

dorikhin, 1999; Slovic, Finucane, Peters, & MacGregor, 2004), and deficits in emotional processing can impair the quality of decision making (e.g., Bechara & Damasio, 2005). The present study was designed to examine how emotional-motivational factors as indexed by electroencephalographic (EEG) asymmetry over the prefrontal cortex (relative activity of the left and right hemispheres) predict purchase decision for national brand and private-label (grocery) products when their price levels were varied. We also examined the factors influencing frontal EEG asymmetry. We think that frontal EEG asymmetry can potentially broaden our view on emotional-motivational processes affecting purchase decision.

Frontal EEG Asymmetry and Approach/ Withdrawal Motivation

According to Davidson's influential approach-withdrawal motivational model of emotion, the left- and right-anterior brain regions are part of two separate neural systems underlying approach and withdrawal motivation, respectively (e.g., Davidson, 1995, 2004). Relatively greater left frontal activity, either as a

Niklas Ravaja, School of Business, Aalto University, Helsinki, Finland, Helsinki Institute for Information Technology (HIIT), Department of Social Research, University of Helsinki, Finland; Outi Somervuori and Mikko Salminen, School of Business, Aalto University.

Correspondence concerning this article should be addressed to Outi Somervuori, Runeberginkatu 22-24, 00100 Helsinki, Finland. E-mail: outi.somervuori@aalto.fi

trait or a state, indicates a propensity to approach or engage a stimulus, while relatively greater right frontal activity indicates a propensity to withdraw or disengage from a stimulus (for reviews, see Coan & Allen, 2004; Davidson, 2003; Demaree, Everhart, Youngstrom, & Harrison, 2005). Source localization of frontal asymmetry in the alpha frequency band (i.e., the index of frontal asymmetry in EEG studies) has indicated that it reflects activity in the dorsal prefrontal cortex (PFC; Pizzagalli, Sherwood, Henriques, & Davidson, 2005). Trait (resting) prefrontal EEG asymmetry (asymmetrical activation of left-frontal vs. right-frontal brain regions) has been shown to predict state-related emotional changes and responses (e.g., affective responses to emotional film clips; Wheeler, Davidson, & Tomarken, 1993) and to be associated with psychopathology or risk for psychopathology (especially depression and anxiety; e.g., Gotlib, Ranganath, & Rosenfeld, 1998; Wiedemann et al., 1999). Increased resting left-lateralized activity has also been associated with a stronger bias to respond to (monetary) reward-related cues (Pizzagalli et al., 2005). Likewise, resting-state hypoactivity in the right lateral PFC has been found to predict higher monetary risk taking (Gianotti et al., 2009) and a lower willingness to punish in the ultimatum game (Knoch, Gianotti, Baumgartner, & Fehr, 2010).

A relationship between emotional states and concomitant changes in frontal EEG asymmetry has also been established; that is, approach-related emotions (e.g., joy and anger) are associated with relatively greater left frontal activation, whereas withdrawal-related emotions (e.g., disgust and fear) are associated with relatively greater right frontal activation (e.g., Coan & Allen, 2003; Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Ekman & Davidson, 1993; Harmon-Jones, Sigelman, Bohlig, & Harmon-Jones, 2003). Davidson, Marshall, Tomarken, and Henriques (2000) have also argued that anterior asymmetry is associated with pregoal attainment emotion elicited while attempting to achieve a goal (e.g., enthusiasm), but not with postgoal attainment emotion (e.g., contentment; cf. the distinction between wanting and liking; see also Tomarken & Zald, 2009). The state engagement in approach-related responses and perceived high as compared with low choice to engage in action (commitment to counterattitudinal or proattitudinal action) has

been shown to increase left-sided frontal activity (Amodio, Devine, & Harmon-Jones, 2007; Harmon-Jones, Harmon-Jones, Serra, & Gable, 2011; see also Harmon-Jones, Lueck, Fearn, & Harmon-Jones, 2006).

Frontal EEG Asymmetry and Purchase Decision

According to Prelec and Loewenstein (1998), a consumer's purchase decision involves a tradeoff between the pleasure derived from consumption and the pain of paying. That is, paying money triggers a perception of loss (i.e., prices are considered as a potential loss), even though it has also been suggested that money spent in buying goods is not "coded" as a loss (no loss in buying hypothesis; Bateman, Kahneman, Munro, Starmer, & Sugden, 2005). In regard to motivational tendencies, anticipatory pleasure of acquisition should be associated with approach motivation, whereas anticipatory pain of paying should be associated with withdrawal motivation. A situation where approach motivation elicited by a preferred product exceeds withdrawal motivation should be associated with an affirmative purchase decision. This (and the aforementioned suggestion that anterior asymmetry is associated with pregoal attainment emotion, but not with postgoal attainment emotion) leads to our first hypothesis:

Hypothesis 1: Relatively greater left frontal activation during the predecision period (i.e., higher alpha asymmetry scores and approach motivation when seeing an image of a product) will predict an affirmative purchase decision, but the decision to purchase the product will not be associated with postdecision alpha asymmetry.

Reference Price and Approach/Withdrawal Motivation

Whereas a price increase from a reference point represents a loss, a price decrease from a reference point represents a gain (e.g., Hardie, Johnson, & Fader, 1993; Putler, 1992). It is also well established that consumers weigh losses from a reference point more heavily than equivalent sized gains, a phenomenon known as loss aversion (Tversky & Kahneman, 1991). Recently, using functional MRI (fMRI), Knutson, Rick, Wimmer, Prelec, and Loewenstein (2007)

found that product preference activated the nucleus accumbens (i.e., a brain region associated with anticipating gain and one of the anatomical structures comprising the approach system; Davidson, 1998), excessive prices activated the right insula (i.e., a region associated with anticipating loss and that is part of a circuit governing positive and negative affect; Davidson, 2004), and reduced prices activated the mesial prefrontal cortex (i.e., a region implicated in integrating gains and losses) prior to the purchase decision. Importantly, activity from each of these regions independently predicted subsequent purchasing decisions. Thus, neural processes underlying purchase decisions may be different depending on whether the price of a product is below or above a reference price. It is also well known that perceptions of quality are positively correlated with price (Rao & Monroe, 1989). Recently, Plassmann, O'Doherty, Shiv, and Rangel (2008) showed that increasing the price of a wine increased subjective reports of flavor pleasantness and activity in medial orbitofrontal cortex (mOFC; i.e., a region thought to encode for experienced pleasantness during experiential tasks; there was also inconclusive evidence that the increase in activation might be more pronounced in the left compared with the right mOFC). This suggests that a high price may elicit conflicting motivational tendencies (i.e., both withdrawal and approach motivation), which may mask the association of asymmetrical frontal cortical activity with a purchase decision. That being so, this association may be more evident when the price of a product is below a reference price. In the present study, we used the normal selling price of a product as a proxy for a reference price (i.e., implicit reference price). Thus, our next hypothesis is:

Hypothesis 2: Relatively greater left frontal activation will be more strongly associated with an affirmative purchase decision when the price of a product is below a normal price (i.e., implicit reference price) compared with when it is above a normal price.

Brand and Approach/Withdrawal Motivation

Evaluative judgments of brands can be based on two distinct types of information or inputs:

(a) declarative information (i.e., brand attributes and brand knowledge) and (b) experiential information (i.e., emotions and experiences evoked by the brand; Brakus, Schmitt, & Zarantonello, 2009; Pham, Cohen, Pracejus, & Hughes, 2001; Schwarz, 2004). Declarative information may be used in a systematic, step-by-step fashion (e.g., expectancy-value model, Fishbein & Ajzen, 1975) or heuristically (e.g., elimination-by-aspects, relational heuristics; for a review, see Bettman & Luce, 1998; see also Maheswaran, Mackie, & Chaiken, 1992). The process where judgments and decisions are based on subjective affective responses to the target, which appear to be seen as indicative of the target's value, has been referred to as the "How-do-I-feel about-it?" heuristic (involving conscious inspection of feelings toward the target; Pham, 1998; Pham et al., 2001) and the "affect heuristic" (encompassing conscious and nonconscious affective influences; Slovic, Finucane, Peters, & MacGregor, 2007). Likewise, the "somatic marker hypothesis," proposed by Damasio and colleagues, suggests that decision process is consciously or nonconsciously influenced by marker signals that arise in bioregulatory processes expressing themselves in emotions and feelings (e.g., Bechara & Damasio, 2005). Through learning and experience, images of options become "marked" by positive and negative feelings linked directly or indirectly to somatic or bodily states.

Brand associations are formed when interacting with the brand (e.g., store visits and actual consumption) and during prior indirect brand exposures (e.g., via brand communications; Esch et al., 2012). Strong (familiar) brands have been suggested to have stronger and more positive brand associations compared with weak (familiar) brands and unfamiliar brands (e.g., Hoeffler & Keller, 2003). Recently, a brain-imaging study by Esch et al. showed that, when evaluating brands, strong brands elicited activations of the pallidum associated with positive emotions, whereas weak and unfamiliar brands elicited activations of the insula associated with negative emotions. In the present study, we focus on national brand and private-label products. Previous research suggests that the influence of deviations from the reference price on purchase behavior may be different for national brand products and private-label products (for the moderating role of quality-tiers in loss aver-

sion, see [Hankuk & Aggarwal, 2003](#)). Consumers tend to perceive brands in the high-quality tier (e.g., national brands) as offering “comfort, security, and value,” whereas brands in the low-quality tier (e.g., private-label brands), offer lower prices but lower quality too ([Hankuk & Aggarwal, 2003](#)). It is also possible that images of private-label products are not marked by strong positive and negative affective feelings; rather, the associations may be neutral. Thus, purchase objectives and psychological processes underlying purchase decision may be different for national brand products and private-label products. Given the discussion above, it would be expected that emotional-motivational factors play a greater role in determining purchase decision for national brand products compared with private-label products. This leads to the following hypothesis:

Hypothesis 3: Relatively greater predecision left frontal activation will be more strongly associated with an affirmative purchase decision for national brand products compared with private label products.

Perceived Need, Product Quality, and Frontal EEG Asymmetry

We also examined the predictors of frontal EEG asymmetry. Hunger and thirst—signals of biological needs—lead to the motivation to get food and water (i.e., appetitive/approach motivation). Likewise, a consumer’s motivation to purchase a product or service is triggered by an expectation that the object of purchase will satisfy his or her perceived biological or other needs. Recently, [Gable and Harmon-Jones \(2008\)](#) showed that self-reported liking for dessert and time since eaten were associated with greater relative left frontal EEG activation during viewing dessert pictures, but not during viewing neutral pictures. Thus, cues signaling potential satisfaction of perceived needs would be expected to elicit approach motivation and relatively greater left frontal activation.

Product attributes, such as perceived quality (i.e., a consumer’s judgment about the overall superiority or excellence of a product; [Zeithaml, 1998](#)), may also exert an influence on approach motivation. As noted above, consumers tend to anticipate that high-quality products will offer “comfort, security, and value” ([Hankuk & Ag-](#)

[garwal, 2003](#)). Thus, images of high-quality products are expected to be marked by positive feelings, thereby eliciting approach motivation. This leads to the following hypothesis:

Hypothesis 4: Higher perceived need for a product and higher perceived product quality will be associated with greater relative left frontal activation during the predecision period (when seeing an image of a product).

Method

Participants

The participants were 33 right-handed healthy business students (14 males and 19 females), who ranged from 20 to 44 years of age ($M = 27.0$). All participants were students who were responsible for their own household’s grocery purchases.

Design

A 7 (Product Category) \times 2 (Brand) \times 15 (Price) within-subjects design was employed.

Seven product categories were selected for the research: detergent, chocolate, coffee, chips, orange juice, chocolate cookies, and toothpaste. For each product category, two products were selected: one national brand product and one store-labeled product (altogether 14 different products). We selected product categories from which two products could be found that are nearly equal in other components except the product wrapping and brand. The selected national brand products were the market leaders of that product category. A corresponding product was selected from the private label category.

The third factor, the price, included 15 different price levels plus one duplicate for the normal price level to control the participant consistency. Altogether, each product was presented 16 times (16 trials). Each product’s normal selling price at a local supermarket was selected as a reference price. The price decrease and corresponding increase levels were 3%, 6%, 10%, 25%, 40%, 60%, and 75%. The behavioral data (the influence of a price decrease and increase on buying behavior) have been reported in [Somervuori and Ravaja \(2011\)](#).

Product Ratings

After the experiment, the participants filled in a questionnaire on the Internet where they rated each of the products on several dimensions. Perceived product quality was rated on a 5-point scale, ranging from 1 (*poor quality*) to 5 (*high quality*). Perceived need for the product was also rated on a 5-point scale, ranging from 1 (*not at all*) to 5 (*very much*).

Procedure

In the laboratory, the participant was first given instructions on the task (i.e., a modified version of the Savings Hold or Purchase [SHOP] task; Knutson et al., 2007) and tested for task comprehension. After the briefing, the participant filled out an informed consent form. Electrodes were then attached, and the participant was seated on a chair. The participant was left alone in the laboratory for a 7-min rest period, followed by the experiment that took, on the average, 52 minutes. The participants received 40 € in cash to spend on products during the experiment. They were asked to imagine them grocery shopping in a local supermarket and having 40 € (their endowment) to spend. All participants were presented with 224 trials in a random order. Each of the 224 trials consisted of the following phases: (a) a fixation cross on a screen presented for 1 s to focus the attention of the participant to the middle of the screen (fixation period), (b) an image of a product with a price shown for 6 s (predecision period), (c) a prompt on the screen to choose either to purchase the product or not by selecting either Y for yes or N for no, and (d) an interstimulus interval varying randomly from 7 to 9 s while the screen was black. The trials were presented using Presentation 10.4 software.

To ensure the participant's engagement in the purchasing task, one trial for each product was randomly selected to count for real (participants were informed about this in the beginning of the experiment). If the participant had chosen to purchase the product in the randomly selected trial, they paid the price shown in the trial from their endowment. In return for their participation, the participants could keep the purchased products and that part of the endowment they had not spent when leaving the experiment. In

addition, the participants were introduced a bonus schema where they were able to gain additional 5 € bonus if they answered "yes" for more than 30% of the trials.

After finishing with all trials, the electrodes were removed, and the participant was debriefed and thanked for participation.

Assessment of EEG

Electrodes mounted in a stretch-Lycra cap (Electrocap; Electro-Cap International, Eaton, OH) were used to record EEG activity from left and right frontal (F3, F4), central (C3, C4), temporal (T7, T8), parietal (P3, P4), and occipital (O1, O2) scalp sites (10–20 International System; Jasper, 1958). The electrodes were referred to linked ears, and the ground lead was located at the left collarbone (e.g., Harmon-Jones & Allen, 1998). Electrode impedances were reduced to less than 5 kΩ. All signals were amplified by a factor of 50,000 with the Psylab EEG8 amplifiers (Contact Precision Instruments, London, United Kingdom). During the data collection, 1-Hz high-pass and 200-Hz low-pass filters were used; a 50-Hz notch filter was also employed. To facilitate artifact detection, ocular movements were recorded with two electrooculogram (EOG) channels. For vertical eye-movements, the electrodes were placed below and above the right eye; for horizontal eye-movements, the electrodes were placed at the outer canthi of the left and right eye. The data collection was controlled by Psylab SAM2 software, and all signals were sampled at a rate of 1,000 Hz.

Data Reduction and Analysis

After the recordings, the EEG data were filtered with 0.5-Hz high-pass and 70-Hz low-pass filters. For each trial, the EEG data were segmented into three 1-s epochs before stimulus (image of a product) onset (seconds 1 to 3) and eight 1-s epochs after stimulus onset (seconds 4 to 11). For artifact removal, all 1-s epochs containing activity outside the range of $-85 \mu\text{V}$ to $+85 \mu\text{V}$, on any of the EEG or EOG channels, were detected and removed from further analyses. For all the remaining 1-s epochs, the power spectra were derived by the fast Fourier transform (FFT) method with a Hanning window (applied to the distal 10% at each end of the epoch). Power values (in μV^2) within the alpha

(8–12 Hz; [Buzsáki, 2006](#)) frequency range were extracted for each 1-s epoch (in alpha asymmetry research, the 8–13 Hz frequency band has also been used; [Allen, Coan, & Nazarian, 2004](#)). Mean power density values were derived for the following periods: (a) baseline (seconds 1 and 2; i.e., two seconds preceding the fixation period), (b) predecision period (seconds 4 to 9), and (c) postdecision period (seconds 10 and 11; for another example of short stimulus periods, i.e., 3-s affective picture viewing, in alpha asymmetry research, see [Harmon-Jones et al., 2006](#)). As in previous research ([Allen et al., 2004](#)), a frontal asymmetry index (natural log of alpha power on the right minus natural log of alpha power on the left) was computed for each period, using midfrontal sites (F3, F4). For comparison purposes, asymmetry indexes for the other sites (C3/4, T7/8, P3/4, O1/2) were also computed. Because cortical alpha power is inversely related to cortical activity ([Cook, O'Hara, Uijtdehaage, Mandelkern, & Leuchter, 1998](#); [Laufs, Kleinschmidt, et al., 2003](#); [Laufs, Krakow, et al., 2003](#); for a review, see [Allen et al., 2004](#)), higher scores on the index indicate greater relative left hemisphere activity. Change scores for alpha asymmetry (Δ alpha asymmetry) were computed by subtracting baseline alpha asymmetry from predecision alpha asymmetry and postdecision alpha asymmetry (cf. [Allen, Harmon-Jones, & Cavender, 2001](#); [Papousek & Schulter, 2002](#)); these change scores reflected changes in asymmetry from the (local) baseline of each trial.

All data were analyzed using the Generalized Estimating Equations (GEE) procedure in SPSS. In the GEE procedure, the dependent variable is linearly related to the factors and covariates via a specified link function. The model allows for the dependent variable to have a non-normal distribution and covers widely used statistical models (e.g., logistic models for binary data). The GEE procedure extends the generalized linear model to allow for analysis of repeated measurements or other correlated observations. The GEE approach requires the specification of the correlation structure of the repeated observations of the dependent variable, distribution of the dependent variable, and link function. The GEE models were introduced by [Liang and Zeger \(1986\)](#), and the method has received wide use in medical and life science research ([Ballinger, 2008](#)).

We specified Participant ID as the subject variable and trial number as the within-subject variable. On the basis of the Quasi-likelihood under Independence Model Criterion (QIC), we specified unstructured as the structure of the working correlation matrix. When predicting purchase decisions, we specified a binomial distribution with logistic link. When predicting EEG alpha asymmetry, we specified a normal distribution with identity as the link function. The terms included in different models are described under the Results section.

Results

[Table 1](#) shows the results of the GEE analyses for purchase decision. The purchase decision was affirmative in 38% of the trials ($n = 224$).

Hypothesis 1

Hypothesis 1 predicted that relatively greater left frontal activation during the predecision period (i.e., higher alpha asymmetry scores and approach motivation when seeing an image of a product) will predict an affirmative purchase decision, but the decision to purchase the product will not be associated with postdecision alpha asymmetry. When testing Hypothesis 1 and Hypothesis 2, product category, normal price, price multiplier, Δ alpha asymmetry, and the Dichotomized Price Multiplier \times Δ Alpha Asymmetry interaction were included in the GEE model. As predicted, the results revealed a significant main effect for predecision Δ alpha asymmetry in predicting purchase decision, $p < .001$. That is, the relatively greater left frontal activation was (i.e., higher approach motivation), the more likely the participant was to purchase a product.

To examine whether the effect was specific to the frontal regions as predicted, we also performed the analysis using the other asymmetry indexes as predictors. Predecision Δ alpha asymmetry in the central and temporal regions was not significantly associated with purchase decision, $ps > .10$. However, the parietal alpha asymmetry index was positively associated, $B = .141$, $SE = .033$, Wald's chi-square ($df = 1$) = 17.85, $p < .001$, and occipital asymmetry index was negatively associated, $B = -.072$, $SE = .030$, Wald's chi-square ($df = 1$) = 5.92, $p =$

Table 1
Results of Generalized Estimating Equations (GEE) Analysis of Purchase Decision Data

Variable	B	SE	Wald χ^2	df	p
Hypothesis 1 (Predecision asymmetry)					
(Intercept)	4.740	.224	449.24	1	<.001
Product category					
Detergent	0.913	.130	49.30	1	<.001
Chocolate	0.065	.107	0.37	1	.546
Chips	0.378	.100	14.14	1	<.001
Coffee	0.459	.112	16.75	1	<.001
Chocolate cookies	0.402	.114	12.45	1	<.001
Orange juice	0.075	.101	0.56	1	.455
Toothpaste	0 ^a				
Normal price	-0.996	.067	219.68	1	<.001
Price multiplier	-3.117	.166	351.10	1	<.001
Δ Alpha asymmetry ^b	0.199	.034	33.62	1	<.001
Dichotomized Price Multiplier \times Δ Alpha Asymmetry ^b	-0.247	.045	30.04	1	<.001
Hypothesis 1 (Postdecision asymmetry)					
(Intercept)	3.790	.162	550.90	1	<.001
Product category					
Detergent	0.761	.086	78.53	1	<.001
Chocolate	0.102	.095	1.16	1	.282
Chips	0.467	.081	32.87	1	<.001
Coffee	0.167	.074	5.13	1	.024
Chocolate cookies	0.675	.091	54.84	1	<.001
Orange juice	-0.035	.086	0.16	1	.687
Toothpaste	0 ^a				
Normal price	-0.963	.036	719.21	1	<.001
Price multiplier	-3.073	.100	949.09	1	<.001
Δ Alpha asymmetry ^b	0.210	.028	56.88	1	<.001
Dichotomized Price Multiplier \times Δ Alpha Asymmetry ^b	-0.322	.035	86.72	1	<.001
Hypothesis 2 (Predecision asymmetry)					
(Intercept)	4.272	.168	649.89	1	<.001
Normal price	-0.701	.038	340.54	1	<.001
Price multiplier	-3.292	.148	498.22	1	<.001
Brand	-0.513	.079	42.04	1	<.001
Δ Alpha asymmetry ^b	0.059	.013	20.52	1	<.001
Brand \times Δ Alpha Asymmetry ^b	0.074	.035	4.50	1	.034
Hypothesis 2 (Postdecision asymmetry)					
(Intercept)	3.167	.133	570.78	1	<.001
Normal price	-0.567	.046	151.10	1	<.001
Price multiplier	-2.947	.084	1231.33	1	<.001
Brand	-0.330	.061	29.39	1	<.001
Δ Alpha asymmetry ^b	0.026	.020	1.73	1	.189
Brand \times Δ Alpha Asymmetry ^b	0.269	.025	113.80	1	<.001

Note. For purchase decision, 0 = not buying (reference category), 1 = buying.

^a Set to zero because this parameter is redundant. ^b Predecision alpha asymmetry ($\ln[F4/F3]$) minus baseline alpha asymmetry.

.015, with an affirmative purchase decision (these associations were weaker compared with that found for frontal asymmetry).

In disagreement with Hypothesis 1, also high postdecision Δ alpha asymmetry scores were significantly related to an affirmative purchase

decision, $p < .001$. In addition to the computation of postdecision Δ alpha asymmetry described above (i.e., postdecision alpha asymmetry minus baseline alpha asymmetry), we computed postdecision Δ alpha asymmetry also by subtracting predecision alpha asymmetry from postdecision

alpha asymmetry (this change score reflected additional changes in asymmetry not accounted for by the predecision period). Also this alternative postdecision Δ alpha asymmetry score was positively related to an affirmative purchase decision, $B = .133$, $SE = .025$, Wald's chi-square ($df = 1$) = 27.67, $p < .001$.

Hypothesis 2

Hypothesis 2 suggested that relatively greater left frontal activation will be more strongly associated with an affirmative purchase decision when the price of a product is below the normal price (i.e., implicit reference price) compared with when it is above the normal price. The results showed that, in addition to the significant main effect for Δ alpha asymmetry, there was a significant Dichotomized Price Multiplier \times (Pre-Decision) Δ Alpha Asymmetry interaction in predicting purchase decision, $p < .001$. That is, as predicted, predecision Δ alpha asymmetry was positively related to an affirmative purchase decision when the price of a product was below the normal price, but not when it was above the normal price (see the top panel of Figure 1). The results revealed also a significant Dichotomized Price Multiplier \times (Post-Decision) Δ Alpha Asymmetry interaction for purchase decision, $p < .001$. As was the case for predecision Δ alpha asymmetry, postdecision Δ alpha asymmetry was associated with an affirmative purchase decision only when the price of a product was below the normal price.

Hypothesis 3

Hypothesis 3 suggested that relatively greater left frontal activation would be more strongly associated with an affirmative purchase decision for national brand products compared with private label products. When testing Hypothesis 3, normal price, price multiplier, brand, Δ alpha asymmetry, and the Brand \times Δ Alpha Asymmetry interaction were included in the GEE model. The results showed that both the Brand \times (Pre-Decision) Δ Alpha Asymmetry interaction and Brand \times (Post-Decision) Δ Alpha Asymmetry interaction were significant in predicting purchase decision, $p = .034$ and $< .001$, respectively. In agreement with Hypothesis 3, predecision and postdecision Δ alpha asymmetry scores were more strongly positively associated with an affirmative purchase decision for na-

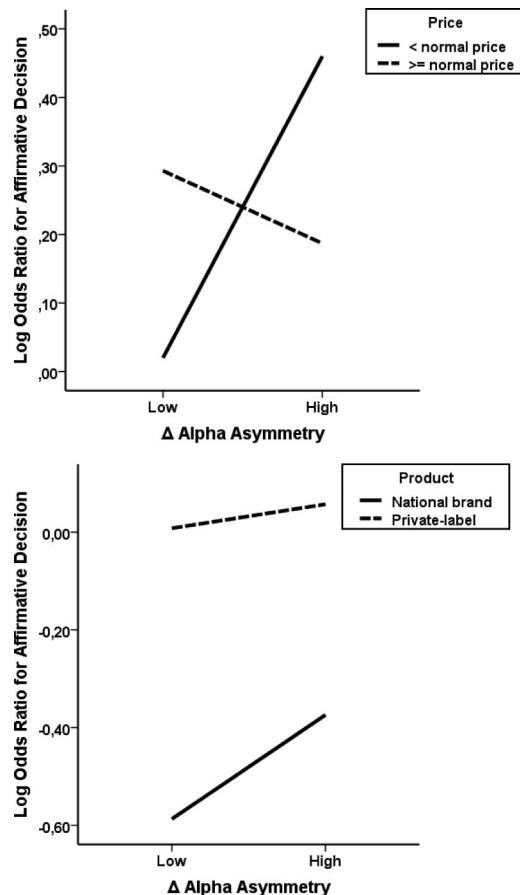


Figure 1. The relationship of predecision Δ alpha asymmetry with purchase decision as a function of dichotomized price multiplier (top panel) and product (national brand or private label; bottom panel; low Δ alpha asymmetry = $M - 1.5$ SD; high Δ alpha asymmetry = $M + 1.5$ SD).

tional brand products compared with private label products (see the bottom panel of Figure 1).

Hypothesis 4

Hypothesis 4 predicted that higher perceived product quality and need would be associated with relatively greater left frontal activation during the predecision period (i.e., higher approach motivation when seeing an image of a product). When testing Hypothesis 4, perceived product quality and need for the product were included in the GEE model. As predicted, the results revealed significant main effects for both perceived product quality, $B = .009$, $SE = .003$, Wald's chi-square ($df = 1$) = 8.80, $p = .003$,

and perceived need, $B = .016$, $SE = .002$, Wald $\chi^2(df = 1) = 45.32$, $p < .001$, in predicting predecision Δ alpha asymmetry. That is, both perceived product quality and perceived need were positively associated with relatively greater left frontal activation.

Discussion

In the present investigation, the authors examined (a) how approach motivation as indexed by EEG asymmetry over the prefrontal cortex predicts purchase decision for national brand and private-label (grocery) products when their prices were varied and (b) the factors influencing frontal EEG asymmetry.

Frontal EEG Asymmetry and Purchase Decision

As hypothesized, we found that relatively greater left frontal activation during the predecision period (i.e., higher approach motivation when seeing an image of a product) predicted an affirmative purchase decision. This is the first study to show that frontal EEG asymmetry predicts purchase decision. The present finding supports the view that a situation where approach motivation evoked by anticipatory pleasure of acquisition exceeds withdrawal motivation evoked by anticipatory pain of paying is associated with an affirmative purchase decision (see Prelec & Loewenstein, 1998). As expected, central and temporal asymmetry indices were not related to purchase decision. However, the parietal asymmetry index was positively associated with an affirmative purchase decision, although most previous studies have found no association of parietal asymmetry with emotional-motivational processes (e.g., Amodio, Shah, Sigelman, Brazy, & Harmon-Jones, 2004; Davidson et al., 1990; but for an exception, see Heller, 1993). Also unexpectedly, relatively greater right occipital activation was related to an affirmative purchase decision. What may account for this association is unknown at the present time (the association was weaker compared with that found for frontal asymmetry and the possibility of Type I error cannot be ruled out).

As opposed to our expectation, we found that also relatively greater postdecision left frontal activation was related to an affirmative purchase

decision (this was the case also for an increase in relative left frontal activation from the predecision to postdecision period). This finding appears to be in disagreement with the suggestion that anterior asymmetry is associated with pregoal attainment emotion, but not with postgoal attainment emotion (Davidson et al., 2000). However, the present study design was not optimal for testing the latter part of our hypothesis, given the procedure that, after completing all trials, only one trial/decision for each product was randomly selected to count for real. That is, at the time of the decision, the participant didn't know whether he or she had really achieved his or her goal (whether an affirmative purchase decision resulted in acquisition of a preferred product). In effect, the fact that both relatively greater predecision and postdecision left frontal activation was related to an affirmative purchase decision increases our confidence in the present findings (Type I error is less likely).

The Moderating Influence of Reference Price and Brand

As also expected, the results showed that greater relative left frontal activation was more strongly related to an affirmative purchase decision when the price of a product was below the normal price (i.e., implicit reference price) compared with when it was above the normal price. This was the case for both predecision and postdecision alpha asymmetry, again increasing our confidence in the finding. Our finding may suggest that there are conflicting motivational tendencies (i.e., both withdrawal and approach motivation) when the price of a product is above the reference price, which may mask the association of frontal EEG asymmetry with a purchase decision. That is, a price increase from a reference point represents a loss (e.g., Hardie et al., 1993; Putler, 1992), which would be expected to elicit withdrawal motivation. However, a high price may also elicit a perception of higher quality, thereby potentially eliciting also approach motivation (cf. Plassmann et al., 2008). The present finding is also in line with the suggestion that neural processes underlying purchase decisions are different depending on whether the price of a product is below or above a reference price (Knutson et al., 2007).

We also found that greater relative left frontal activation was more strongly associated with an affirmative purchase decision for national brand

products compared with private-label products. Again, this was the case for both predecision and postdecision alpha asymmetry. This finding suggests that emotional-motivational factors play a greater role in determining purchase decision for national brand products compared with private-label products. Brand associations have previously been suggested as being stronger and more positive for strong (familiar) brands compared with weak (familiar) brands and unfamiliar brands (e.g., Hoeffler & Keller, 2003). The present results suggest that images of private-label products may not be marked by strong positive or negative affective feelings (see, e.g., Bechara & Damasio, 2005); the brand associations for private-label products may rather be neutral. Apparently, not only purchase objectives but also psychological processes underlying purchase decision are different for national brand products and private-label products.

Predictors of Frontal EEG Asymmetry

In agreement with our hypothesis, the results showed that higher perceived need for the product was associated with greater relative left frontal EEG activation during the predecision period (when seeing an image of a product). Given that a need elicits appetitive/approach motivation, this finding supports the validity of frontal EEG asymmetry as a measure of approach motivation. The present finding is in line with the view that a consumer's motivation to purchase a product or service is triggered by an expectation that the object of purchase will satisfy his or her perceived needs. It is also in line with prior research showing that time since eaten (indexing a biological need) was associated with greater relative left frontal EEG activation during viewing dessert pictures, but not during viewing neutral pictures (Gable & Harmon-Jones, 2008). We also found that higher perceived product quality was related to greater relative left frontal activation during the predecision period. Given that high-quality products are anticipated as offering "comfort, security, and value" (Hankuk & Aggarwal, 2003), their images are expected to be marked by positive feelings, thereby eliciting approach motivation.

Limitations

Although the modified version of the SHOP task used in the present study entails the advantage

that the decisions made by the participants have real monetary consequences for them, an apparent limitation was that the decision making situation, nevertheless, differs from that typical for purchasing grocery products (the present situation resembles, to some extent, a Web auction). It is unclear, however, whether this difference should have any influence on the results obtained. It should also be noted that the present results apply to grocery products of relatively low price. One may expect, however, that the results would have been even stronger for more expensive products.

An additional limitation relates a procedural issue that produces interpretational difficulties. As has been customary in most of the previous research, we quantified asymmetry as the difference between right frontal activation and left frontal activation. This computation of asymmetry implies that there is a single bipolar (reciprocal) continuum of cortical activation, thereby being in contrast with the view that approach and withdrawal motivation are largely independent (Ito & Cacioppo, 1999). It is also of note that several different patterns of activation may be represented by the same asymmetry score (e.g., a moderate asymmetry score can indicate either high left and high right frontal activation or low left and low right frontal activation). Also, the limited spatial resolution provided by EEG is an issue that arises in regional EEG research (and we used only a few electrodes).

Finally, it might have been advantageous to have a separate product period (image of a product without a price) and a price period (image of a product with a price) in the trials of the experiment (see Knutson et al., 2007). This would have been optimal for studying separately the approach/withdrawal motivation elicited by preferred products and prices, although the present factorial design varying the product and price is basically also able to tease out this information.

Conclusions

The present study showed that greater relative left frontal EEG activation during the predecision period predicted an affirmative purchase decision for grocery products. This relationship was stronger when the price of a product was below a normal price (implicit reference price) compared with when it was above

a normal price, suggesting that there may be conflicting motivational tendencies (i.e., both withdrawal and approach motivation) when the price of a product is above the reference price. The results also suggested that emotional-motivational factors play a greater role in determining purchase decision for national brand products (the images of which are marked by strong affective feelings) compared with private-label products. In general, the results provide further evidence for the importance of emotional-motivational factors in purchase decision. This study also supports the usefulness of frontal EEG asymmetry as a measure of approach/withdrawal motivation when studying purchase decision. Frontal EEG asymmetry adds a new dimension to our understanding of emotional-motivational processes affecting purchase decision—a dimension that we cannot necessarily tap, if we only record behavioral responses.

References

- Allen, J. J. B., Coan, J. A., & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological Psychology*, 67, 183–218. doi: [10.1016/j.biopsych.2004.03.007](https://doi.org/10.1016/j.biopsych.2004.03.007)
- Allen, J. J. B., Harmon-Jones, E., & Cavender, J. H. (2001). Manipulation of frontal EEG asymmetry through biofeedback alters self-reported emotional responses and facial EMG. *Psychophysiology*, 38, 685–693. doi: [10.1111/1469-8986.3840685](https://doi.org/10.1111/1469-8986.3840685)
- Amodio, D. M., Devine, P. G., & Harmon-Jones, E. (2007). A dynamic model of guilt: Implications for motivation and self-regulation in the context of prejudice. *Psychological Science*, 18, 524–530. doi: [10.1111/j.1467-9280.2007.01933.x](https://doi.org/10.1111/j.1467-9280.2007.01933.x)
- Amodio, D. M., Shah, J. Y., Sigelman, J., Brazy, P. C., & Harmon-Jones, E. (2004). Implicit regulatory focus associated with asymmetrical frontal cortical activity. *Journal of Experimental Social Psychology*, 40, 225–232. doi: [10.1016/S0022-1031\(03\)00100-8](https://doi.org/10.1016/S0022-1031(03)00100-8)
- Ballinger, G. A. (2008). Models for intensive longitudinal data. *Organizational Research Methods*, 11, 412–414. doi: [10.1177/1094428107309307](https://doi.org/10.1177/1094428107309307)
- Bateman, I., Kahneman, D., Munro, A., Starmer, C., & Sugden, R. (2005). Testing competing models of loss aversion: An adversarial collaboration. *Journal of Public Economics*, 89, 1561–1580. doi: [10.1016/j.jpubeco.2004.06.013](https://doi.org/10.1016/j.jpubeco.2004.06.013)
- Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, 52, 336–372. doi: [10.1016/j.geb.2004.06.010](https://doi.org/10.1016/j.geb.2004.06.010)
- Bernheim, B. D., & Rangel, A. (2004). Addiction and cue-triggered decision processes. *American Economic Review*, 94, 1558–1590. doi: [10.1257/0002828043052222](https://doi.org/10.1257/0002828043052222)
- Bettman, J. R., & Luce, M. F. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, 25, 187–217. doi: [10.1086/209535](https://doi.org/10.1086/209535)
- Brakus, J. J., Schmitt, B. H., & Zarantonello, L. (2009). Brand experience: What is it? How is it measured? Does it affect loyalty? *Journal of Marketing*, 73, 52–68. doi: [10.1509/jmkg.73.3.52](https://doi.org/10.1509/jmkg.73.3.52)
- Buzsáki, G. (2006). *Rhythms of the brain*. New York, NY: Oxford University Press. doi: [10.1093/acprof:oso/9780195301069.001.0001](https://doi.org/10.1093/acprof:oso/9780195301069.001.0001)
- Coan, J. A., & Allen, J. J. B. (2003). Varieties of emotional experience during voluntary emotional facial expressions. *Annals of the New York Academy of Sciences*, 1000, 375–379. doi: [10.11196/annals.1280.034](https://doi.org/10.11196/annals.1280.034)
- Coan, J. A., & Allen, J. J. B. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological Psychology*, 67, 7–49. doi: [10.1016/j.biopsych.2004.03.002](https://doi.org/10.1016/j.biopsych.2004.03.002)
- Cook, I. A., O'Hara, R., Uijtdehaage, S. H. J., Mandelkern, M., & Leuchter, A. F. (1998). Assessing the accuracy of topographic EEG mapping for determining local brain function. *Electroencephalography and Clinical Neurophysiology*, 107, 408–414. doi: [10.1016/S0013-4694\(98\)00092-3](https://doi.org/10.1016/S0013-4694(98)00092-3)
- Davidson, R. J. (1995). Cerebral asymmetry, emotion, and affective style. In R. J. Davidson & K. Hugdahl (Eds.), *Brain asymmetry* (pp. 361–387). Cambridge, MA: MIT Press.
- Davidson, R. J. (1998). Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums. *Psychophysiology*, 35, 607–614. doi: [10.1017/S0048577298000134](https://doi.org/10.1017/S0048577298000134)
- Davidson, R. J. (2003). Affective neuroscience and psychophysiology: Toward a synthesis. *Psychophysiology*, 40, 655–665. doi: [10.1111/1469-8986.00067](https://doi.org/10.1111/1469-8986.00067)
- Davidson, R. J. (2004). What does the prefrontal cortex “do” in affect: Perspectives on frontal EEG asymmetry research. *Biological Psychology*, 67, 219–233. doi: [10.1016/j.biopsych.2004.03.008](https://doi.org/10.1016/j.biopsych.2004.03.008)
- Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I. *Journal of Personality and Social Psychology*, 58, 330–341. doi: [10.1037/0022-3514.58.2.330](https://doi.org/10.1037/0022-3514.58.2.330)
- Davidson, R. J., Marshall, J. R., Tomarken, A. J., & Henriques, J. B. (2000). While a phobic waits: Regional brain electrical and autonomic activity in social phobics during anticipation of public speak-

- ing. *Biological Psychiatry*, 47, 85–95. doi: [10.1016/S0006-3223\(99\)00222-X](https://doi.org/10.1016/S0006-3223(99)00222-X)
- Demaree, H. A., Everhart, D. E., Youngstrom, E. A., & Harrison, D. W. (2005). Brain lateralization of emotional processing: Historical roots and a future incorporating “dominance”. *Behavioral and Cognitive Neuroscience Reviews*, 4, 3–20. doi: [10.1177/1534582305276837](https://doi.org/10.1177/1534582305276837)
- Ekman, P., & Davidson, R. J. (1993). Voluntary smiling changes regional brain activity. *Psychological Science*, 4, 342–345. doi: [10.1111/j.1467-9280.1993.tb00576.x](https://doi.org/10.1111/j.1467-9280.1993.tb00576.x)
- Esch, F.-R., Möll, T., Schmitt, B., Elger, C. E., Neuhaus, C., & Weber, B. (2012). Brands on the brain: Do consumers use declarative information or experienced emotions to evaluate brands? *Journal of Consumer Psychology*, 22, 75–85. doi: [10.1016/j.jcps.2010.08.004](https://doi.org/10.1016/j.jcps.2010.08.004)
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Gable, P., & Harmon-Jones, E. (2008). Relative left frontal activation to appetitive stimuli: Considering the role of individual differences. *Psychophysiology*, 45, 275–278. doi: [10.1111/j.1469-8986.2007.00627.x](https://doi.org/10.1111/j.1469-8986.2007.00627.x)
- Gianotti, L. R. R., Knoch, D., Faber, P. L., Lehmann, D., Pascual-Marqui, R. D., Diez, C., . . . Fehr, E. (2009). Tonic activity level in the right prefrontal cortex predicts individuals' risk taking. *Psychological Science*, 20, 33–38. doi: [10.1111/j.1467-9280.2008.02260.x](https://doi.org/10.1111/j.1467-9280.2008.02260.x)
- Gotlib, I. A., Ranganath, C., & Rosenfeld, J. P. (1998). Frontal EEG alpha asymmetry, depression, and cognitive functioning. *Cognition and Emotion*, 12, 449–478. doi: [10.1080/02699398379673](https://doi.org/10.1080/02699398379673)
- Hankuk, T. C., & Aggarwal, P. (2003). When gains exceed losses: Attribute trade-offs and prospect theory. *Advances in Consumer Research*, 30, 118–124.
- Hardie, B. G. S., Johnson, E. J., & Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science*, 12, 378–394. doi: [10.1287/mksc.12.4.378](https://doi.org/10.1287/mksc.12.4.378)
- Harmon-Jones, E., & Allen, J. J. B. (1998). Anger and frontal brain activity: EEG asymmetry consistent with approach motivation despite negative affective valence. *Journal of Personality and Social Psychology*, 74, 1310–1316. doi: [10.1037/0022-3514.74.5.1310](https://doi.org/10.1037/0022-3514.74.5.1310)
- Harmon-Jones, E., Harmon-Jones, C., Serra, R., & Gable, P. A. (2011). The effect of commitment on relative left frontal cortical activity: Tests of the action-based model of dissonance. *Personality and Social Psychology Bulletin*, 37, 395–408. doi: [10.1177/0146167210397059](https://doi.org/10.1177/0146167210397059)
- Harmon-Jones, E., Lueck, L., Fearn, M., & Harmon-Jones, C. (2006). The effect of personal relevance and approach-related action expectation on relative left frontal cortical activity. *Psychological Science*, 17, 434–440. doi: [10.1111/j.1467-9280.2006.01724.x](https://doi.org/10.1111/j.1467-9280.2006.01724.x)
- Harmon-Jones, E., Sigelman, J. D., Bohlig, A., & Harmon-Jones, C. (2003). Anger, coping, and frontal cortical activity: The effect of coping potential on anger-induced left frontal activity. *Cognition and Emotion*, 17, 1–24. doi: [10.1080/0269930302278](https://doi.org/10.1080/0269930302278)
- Heller, W. (1993). Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology*, 7, 476–489. doi: [10.1037/0894-4105.7.4.476](https://doi.org/10.1037/0894-4105.7.4.476)
- Hoefller, S., & Keller, K. (2003). The marketing advantages of strong brands. *Journal of Brand Management*, 10, 421–445. doi: [10.1057/palgrave.bm.2540139](https://doi.org/10.1057/palgrave.bm.2540139)
- Ito, T., & Cacioppo, J. (1999). The psychophysiology of utility appraisals. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Wellbeing: The foundations of hedonic psychology* (pp. 470–488). New York, NY: Russell Sage Foundation.
- Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 10, 371–375.
- Kahneman, D., Ritov, I., & Schkade, D. (1999). Economic preferences or attitude expressions? An analysis of dollar responses to public issues. *Journal of Risk & Uncertainty*, 19, 203–235. doi: [10.1023/A:1007835629236](https://doi.org/10.1023/A:1007835629236)
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291. doi: [10.2307/1914185](https://doi.org/10.2307/1914185)
- Knoch, D., Gianotti, L. R. R., Baumgartner, T., & Fehr, E. (2010). A neural marker of costly punishment behavior. *Psychological Science*, 21, 337–342. doi: [10.1177/0956797609360750](https://doi.org/10.1177/0956797609360750)
- Knutson, B., Rick, S., Wimmer, E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, 53, 147–156. doi: [10.1016/j.neuron.2006.11.010](https://doi.org/10.1016/j.neuron.2006.11.010)
- Laufs, H., Kleinschmidt, A., Beyerle, A., Eger, E., Salek-Haddadi, A., Preibisch, C., & Krakow, K. (2003). EEG-correlated fMRI of human alpha activity. *Neuroimage*, 19, 1463–1476. doi: [10.1016/S1053-8119\(03\)00286-6](https://doi.org/10.1016/S1053-8119(03)00286-6)
- Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). Electrocorticographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 11053–11058. doi: [10.1073/pnas.1831638100](https://doi.org/10.1073/pnas.1831638100)
- Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data-analysis using generalized linear-models. *Bi-*

- ometrika*, 73, 1311053–22. doi:10.1093/biomet/73.1.13
- Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 619–642). Oxford, UK: Oxford University Press.
- Maheswaran, D., Mackie, D. M., & Chaiken, S. (1992). Brand name as a heuristic cue: The Effects of task importance and expectancy confirmation on consumer judgments. *Journal of Consumer Psychology*, 1, 317–336. doi:10.1016/S1057-7408(08)80058-7
- Naqvi, N. H., Shiv, B., & Bechara, A. (2006). The role of emotion in decision-making: A cognitive neuroscience perspective. *Current Directions in Psychological Science*, 15, 260–264. doi:10.1111/j.1467-8721.2006.00448.x
- Oullier, O., Kirman, A. P., & Kelso, J. A. S. (2008). The coordination dynamics of economic decision making: A multilevel approach to social neuroeconomics. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16, 557–571. doi:10.1109/TNSRE.2008.2009960
- Papousek, I., & Schulter, G. (2002). Covariations of EEG asymmetries and emotional states indicate that activity at frontopolar locations is particularly affected by state factors. *Psychophysiology*, 39, 350–360. doi:10.1017/S0048577201393083
- Pham, M. T. (1998). Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, 25, 144–159. doi:10.1086/209532
- Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research*, 28, 167–189. doi:10.1086/322896
- Pizzagalli, D. A., Sherwood, R. J., Henriques, J. B., & Davidson, R. J. (2005). Frontal brain asymmetry and reward responsiveness: A source-localization study. *Psychological Science*, 16, 805–813. doi:10.1111/j.1467-9280.2005.01618.x
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences*, 105, 1050–1054. doi:10.1073/pnas.0706929105
- Prelec, D., & Loewenstein, G. F. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science*, 17, 4–28. doi:10.1287/mksc.17.1.4
- Putler, D. S. (1992). Incorporating reference price effects into a theory of consumer choice. *Market-*
- ing Science*, 11, 287–309. doi:10.1287/mksc.11.3.287
- Rao, A., & Monroe, K. B. (1989). The effect of price, brand name, and store name on buyers' perceptions of product quality: An integrative review. *Journal of Marketing Research*, 26, 351–357. doi:10.2307/3172907
- Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision-making. *Journal of Consumer Psychology*, 14, 332–348. doi:10.1207/s15327663jcp1404_2
- Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26, 278–292. doi:10.1086/209563
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, 24, 311–322. doi:10.1111/j.0272-4332.2004.00433.x
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research*, 177, 1333–1352. doi:10.1016/j.ejor.2005.04.006
- Somervuori, O., & Ravaja, N. (2011). Purchase behavior and psychophysiological responses when prices are increased and decreased. Manuscript submitted for publication.
- Tomarken, A. J., & Zald, D. H. (2009). Conceptual, methodological, and empirical ambiguities in the linkage between anger and approach: Comment on Carver and Harmon-Jones (2009). *Psychological Bulletin*, 135, 209–214. doi:10.1037/a0014735
- Tversky, A., & Kahneman, D. (1991). Loss aversion and riskless choice: A reference dependent model. *Quarterly Journal of Economics*, 106, 1039–1061. doi:10.2307/2937956
- Wheeler, R. E., Davidson, R. J., & Tomarken, A. J. (1993). Frontal brain asymmetry and emotional reactivity: A biological substrate of affective style. *Psychophysiology*, 30, 82–89. doi:10.1111/j.1469-8986.1993.tb03207.x
- Wiedemann, G., Pauli, P., Dengler, W., Lutzenberger, W., Birbaumer, N., & Buchkremer, G. (1999). Frontal brain asymmetry as a biological substrate of emotions in patients with panic disorders. *Archives of General Psychiatry*, 56, 78–84. doi:10.1001/archpsyc.56.1.78
- Zeithaml, V. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing*, 52, 278–22. doi:10.2307/1251446

Received February 6, 2012
 Revision received July 17, 2012
 Accepted August 6, 2012 ■

and Ipl1 moves from kinetochores to spindle microtubules shortly after the initiation of anaphase (5, 27, 28). Microtubule attachment to kinetochores in anaphase may be stabilized by the loss of Ipl1, helping to keep the checkpoint inactive. However, Ipl1 mutants respond to treatment with nocodazole, whereas anaphase-arrested cells do not (Fig. 1A), which suggests that additional factors, such as Mps1 degradation, have turned off the checkpoint in anaphase (29). The organization of other “chromosomal passenger proteins” also changes as cells enter anaphase (30), as do spindle microtubule dynamics (31), and these factors may also influence checkpoint behavior in anaphase. Finally, the checkpoint destabilizes Cdc20, as well as inhibits its activity, which reinforces the mutual antagonism between the checkpoint and APC^{Cdc20}.

We have presented evidence for a mechanism that inactivates the spindle checkpoint as yeast cells enter anaphase. When mitosis starts, the APC is off, the checkpoint is on, and checkpoint proteins are stable. As long as one chromosome has not aligned, the checkpoint inhibits the APC. When this chromosome biorients, a threshold is crossed, the APC becomes active, cells enter anaphase, and the destruction of Mps1 (and possibly other checkpoint proteins) permanently inactivates the checkpoint. The opposing activities of the checkpoint and the APC let cells

switch rapidly between prometaphase, when they can sensitively monitor chromosome alignment, and anaphase, when they are irreversibly committed to entering the next cell cycle, despite the lack of tension at the kinetochores.

References and Notes

- K. Nasmyth, *Science* **297**, 559 (2002).
- R. B. Nicklas, C. A. Koch, *J. Cell Biol.* **43**, 40 (1969).
- M. A. Hoyt, L. Totis, B. T. Roberts, *Cell* **66**, 507 (1991).
- R. Li, A. W. Murray, *Cell* **66**, 519 (1991).
- B. A. Pinsky, C. Kung, K. M. Shokat, S. Biggins, *Nat. Cell Biol.* **8**, 78 (2006).
- C. L. Rieder, A. Schultz, R. Cole, G. Sluder, *J. Cell Biol.* **127**, 1301 (1994).
- X. Li, R. B. Nicklas, *Nature* **373**, 630 (1995).
- W. A. Wells, A. W. Murray, *J. Cell Biol.* **133**, 75 (1996).
- A. Musacchio, K. G. Hardwick, *Nat. Rev. Mol. Cell Biol.* **3**, 731 (2002).
- L. H. Hwang *et al.*, *Science* **279**, 1041 (1998).
- S. L. Jaspersen, J. F. Charles, R. L. Tinker-Kulberg, D. O. Morgan, *Mol. Biol. Cell* **9**, 2803 (1998).
- O. Cohen-Fix, J. M. Peters, M. W. Kirschner, D. Koshland, *Genes Dev.* **10**, 3081 (1996).
- M. Shirayama, A. Toth, M. Galova, K. Nasmyth, *Nature* **402**, 203 (1999).
- O. Cohen-Fix, D. Koshland, *Genes Dev.* **13**, 1950 (1999).
- F. M. Yeong, H. H. Lim, C. G. Padmashree, U. Surana, *Mol. Cell* **5**, 501 (2000).
- E. Weiss, M. Winey, *J. Cell Biol.* **132**, 111 (1996).
- K. G. Hardwick, E. Weiss, F. C. Luca, M. Winey, A. W. Murray, *Science* **273**, 953 (1996).
- C. A. Ball *et al.*, *Nucleic Acids Res.* **28**, 77 (2000).
- K. G. Hardwick, A. W. Murray, *J. Cell Biol.* **131**, 709 (1995).
- R. Li, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 4989 (1999).
- M. Glotzer, A. W. Murray, M. W. Kirschner, *Nature* **349**, 132 (1991).

22. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; and X, any amino acid.

23. S. L. Jaspersen *et al.*, *Dev. Cell* **7**, 263 (2004).

24. See supporting data on *Science Online*.

25. M. H. Jones *et al.*, *Curr. Biol.* **15**, 160 (2005).

26. J. Pan, R. H. Chen, *Genes Dev.* **18**, 1439 (2004).

27. S. Buvelot, S. Y. Tatsutani, D. Vermaak, S. Biggins, *J. Cell Biol.* **160**, 329 (2003).

28. G. Pereira, E. Schiebel, *Science* **302**, 2120 (2003).

29. S. Biggins, N. Bhalla, A. Chang, D. L. Smith, A. W. Murray, *Genetics* **159**, 453 (2001).

30. P. Vagnarelli, W. C. Earnshaw, *Chromosoma* **113**, 211 (2004).

31. T. Higuchi, F. Uhlmann, *Nature* **433**, 171 (2005).

32. O. Cohen-Fix, D. Koshland, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 14361 (1997).

33. We thank S. C. Schuyler and B. Stern for their support and technical assistance, the members of the Murray and Winey labs for their helpful comments and suggestions, K. Hardwick for providing antibodies, and A. Wenger for his efforts. Supported by NIH grants GM43987 (A.W.M.) and GM51312 (M.W.). S.L.J. is a Leukemia and Lymphoma Society Special Fellow.

Supporting Online Material

www.science.org/cgi/content/full/1127205/DC1

SOM Text

Figs. S1 to S6

Table S1

9 March 2006; accepted 30 May 2006

Published online 6 July 2006;

10.1126/science.1127205

Include this information when citing this paper.

Frames, Biases, and Rational Decision-Making in the Human Brain

Benedetto De Martino,* Dharshan Kumaran, Ben Seymour, Raymond J. Dolan

Human choices are remarkably susceptible to the manner in which options are presented. This so-called “framing effect” represents a striking violation of standard economic accounts of human rationality, although its underlying neurobiology is not understood. We found that the framing effect was specifically associated with amygdala activity, suggesting a key role for an emotional system in mediating decision biases. Moreover, across individuals, orbital and medial prefrontal cortex activity predicted a reduced susceptibility to the framing effect. This finding highlights the importance of incorporating emotional processes within models of human choice and suggests how the brain may modulate the effect of these biasing influences to approximate rationality.

A central tenet of rational decision-making is logical consistency across decisions, regardless of the manner in which available choices are presented. This assumption, known as “extensionality” (1) or “invariance” (2), is a fundamental axiom of game theory (3). However, the proposition that human decisions are “description-invariant” is challenged by a wealth of empirical data (4, 5). Kahneman and Tversky originally described this deviation from

rational decision-making, which they termed the “framing effect,” as a key aspect of prospect theory (6, 7).

Theories of decision-making have tended to emphasize the operation of analytic processes in guiding choice behavior. However, more intuitive or emotional responses can play a key role in human decision-making (8–10). Thus, when taking decisions under conditions when available information is incomplete or overly complex, subjects rely on a number of simplifying heuristics, or efficient rules of thumb, rather than extensive algorithmic processing (11). One suggestion is that the framing effect results from systematic biases in choice behavior arising from an affect heuristic under-

written by an emotional system (12, 13). However, despite the substantial role of the framing effect in influencing human decision-making, the underlying neurobiological basis is not understood.

We investigated the neurobiological basis of the framing effect by means of functional magnetic resonance imaging (fMRI) and a novel financial decision-making task. Participants (20 university students or graduates) received a message indicating the amount of money that they would initially receive in that trial (e.g., “You receive £50”). Subjects then had to choose between a “sure” option and a “gamble” option presented in the context of two different frames. The “sure” option was formulated as either the amount of money retained from the initial starting amount (e.g., keep £20 of the £50; “Gain” frame) or as the amount of money lost from the initial amount (e.g., lose £30 of the £50; “Loss” frame). The “gamble” option was identical in both frames and was represented as a pie chart depicting the probability of winning or losing (Fig. 1) (14).

The behavioral results indicated that subjects’ decisions were significantly affected by our framing manipulation, with a marked difference in choices between the two frames (Fig. 2A). Specifically, and in accordance with predictions arising from prospect theory, subjects were risk-averse in the Gain frame, tending to choose the sure option over the gamble option

*To whom correspondence should be addressed. E-mail: b.martino@filion.ac.uk

[gambling on 42.9% of trials; significantly different from 50% ($P < 0.05$, $t_{19} = 1.96$)], and were risk-seeking in the Loss frame, preferring the gamble option [gambling on 61.6% of trials; significantly different from 50% ($P < 0.005$, $t_{19} = 3.31$)]. This effect of frame was consistently expressed across different probabilities and starting amounts (fig. S1).

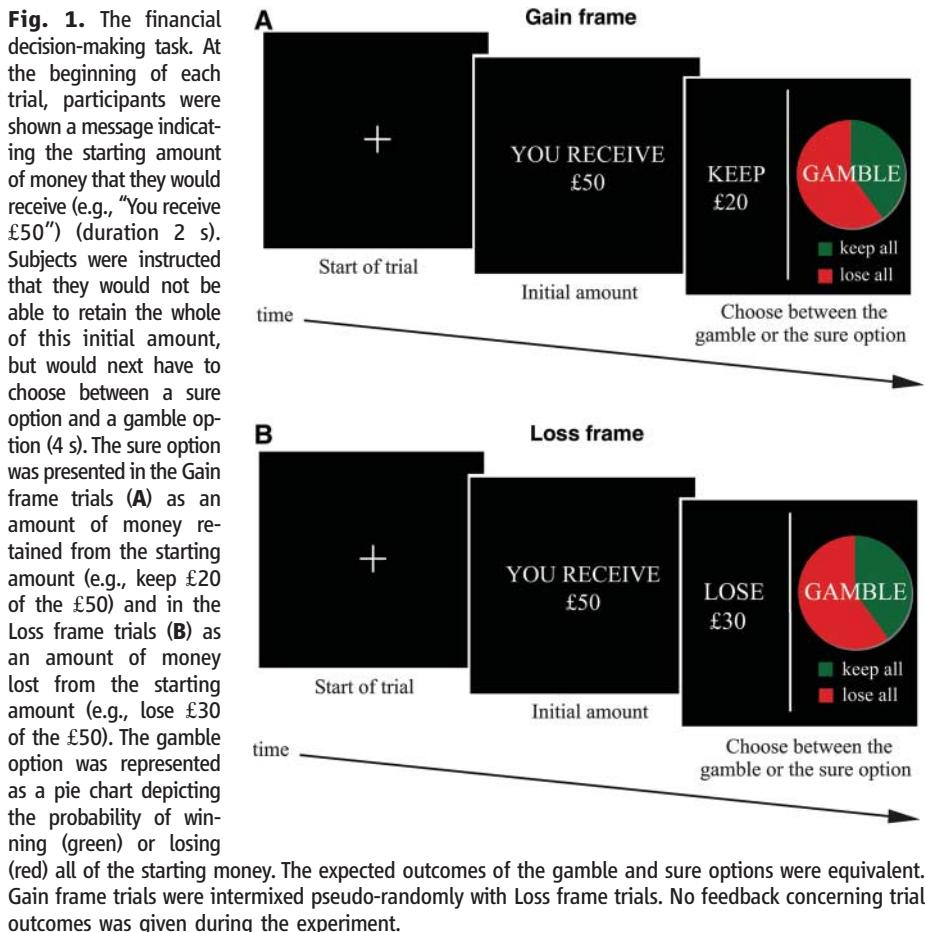


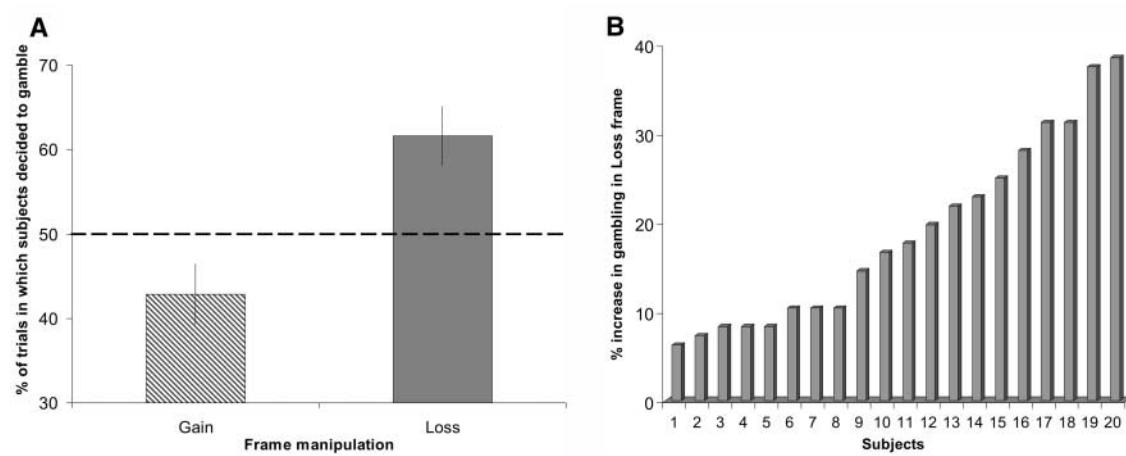
Fig. 2. Behavioral results. (A) Percentages of trials in which subjects chose the gamble option in the Gain frame and the Loss frame. Subjects showed a significant increase in the percentage of trials in which the gamble option was chosen in the Loss frame with respect to the Gain frame [$61.6\% > 42.9\%$ ($P < 0.001$, $t_{19} = 8.06$)]. The dashed line represents risk-neutral behavior (choosing the gamble option in 50% of trials). Error bars denote SEM. (B) Each bar represents, for each individual subject, the percentage difference between how often subjects chose the gamble option in the Loss frame as compared to the Gain frame. A hypothetical value of zero represents a complete indifference to the framing manipulation (i.e., fully "rational" behavior). All participants, to varying degrees, showed an effect of the framing manipulation.

Reaction times for decisions were not affected by frame [Gain frame, 1895 ms; Loss frame, 1884 ms ($P > 0.1$)]; this result provides evidence that difficulty was well matched between the two frames. Moreover, subjects performed highly accurately on "catch" trials (14) (fig. S2) where the expected outcomes of the sure and gamble options were unbalanced,

indicating their continued engagement with the task throughout the experiment. Despite the marked though variable impact of the frame on subjects' choice behavior (Fig. 2B), the majority (16/20) of subjects seemed unaware of any biasing effect when specifically questioned in a debriefing session that followed the experiment.

Subjects performed the behavioral task inside an fMRI scanner, allowing us to obtain continuous measures of regional brain activity. The subjects' individual decisions during the entire fMRI experiment were recorded and used to construct four regressors of interest: sure decisions in the Gain frame (G_{sure}), gamble decisions in the Gain frame (G_{gamble}), sure decisions in the Loss frame (L_{sure}), and gamble decisions in the Loss frame (L_{gamble}).

Given that the frame effect relates to subjects' asymmetrical pattern of decisions across frames, the key experimental contrast of interest is the interaction between the decision to gamble (or not) and the valence of the frame: $[(G_{\text{sure}} + L_{\text{gamble}}) - (G_{\text{gamble}} + L_{\text{sure}})]$. It is noteworthy that this interaction contrast is balanced with respect to both decision type and frame valence. Consequently, we could identify brain areas that were more active when subjects chose in accordance with the frame effect (i.e., $G_{\text{sure}} + L_{\text{gamble}}$), as opposed to when their decisions ran counter to their general behavioral tendency ($G_{\text{gamble}} + L_{\text{sure}}$). This contrast revealed significant activation in the bilateral amygdala (Fig. 3, A and B). To ensure that this activation in the amygdala was not being driven by a significant effect in one frame alone (e.g., Loss frame), we conducted an independent analysis for each frame. This confirmed that robust activation in the amygdala was equally observed for simple effects of decision type (sure or gamble) in each frame separately. Thus, amygdala activation was



significantly greater when subjects decided to choose the sure option in the Gain frame [G_{sure} – G_{gamble}] [Montreal Neurological Institute (MNI)

space coordinates (x, y, z) 18, –4, –24; Z score = 4.0], and the gamble option in the Loss frame [$L_{\text{gamble}} - L_{\text{sure}}$] [MNI space coordinates –16, 0,

–26; Z score = 3.80; 12, 2, –22; Z score = 4.67], in keeping with a central role in mediating the frame effect.

A different pattern of brain activation was identified when subjects made decisions that ran counter to their general behavioral tendency. In this reverse interaction contrast [$(G_{\text{gamble}} + L_{\text{sure}}) - (G_{\text{sure}} + L_{\text{gamble}})$], we observed enhanced activity in the anterior cingulate cortex (ACC) (Fig. 3, C and D) (and to a lesser extent in the bilateral dorsolateral prefrontal cortex at an uncorrected threshold of $P < 0.005$; fig. S3) when subjects chose the gamble option in the Gain frame and the sure option in the Loss frame.

In light of the substantial intersubject variability in behavioral susceptibility to the frame, we next identified subject-specific differences in neural activity associated with their decision bias (that is, the decision \times frame interaction) (Fig. 2A). Using the overall susceptibility of each subject to the frame manipulation as a between-subjects statistical regressor, operationalized as a "rationality index" (14), we found a significant correlation between decreased susceptibility to the framing effect and enhanced activity in the orbital and medial prefrontal cortex (OMPFC), specifically in the right orbitofrontal cortex (R-OFC; $r = 0.8$, $P < 0.001$) and the ventromedial prefrontal cortex (VMPFC; $r = 0.75$, $P < 0.001$) (Fig. 4). In summary, those subjects who acted more rationally exhibited greater activation in OMPFC associated with the frame effect.

Our data provide a neurobiological account of the framing effect, both within and across individuals. Increased activation in the amygdala was associated with subjects' tendency to be risk-averse in the Gain frame and risk-seeking in the Loss frame, supporting the hypothesis that the framing effect is driven by an affect heuristic underwritten by an emotional system. The amygdala plays a key role in value-related prediction and learning, both for negative (aversive) and positive (appetitive) outcomes (15–17). Furthermore, in simple instrumental decision-making tasks in animals, the amygdala appears to mediate decision biases that come from value-related predictions (18). In humans, the amygdala is also implicated in the detection of emotionally relevant information present in contextual and social emotional cues (19). It was previously shown that activation in the amygdala during the passive viewing of surprised faces is significantly modulated by the valence of preceding verbal contextual information (20). Our data extend the role of the amygdala to include processing the type of contextual positive or negative emotional information communicated by the frame in the context of a decision-making task.

In our study, activation of the amygdala was driven by the combination of a subject's decision and the frame in which it took place, rather than by the valence of the frame per se. Consequently, our findings indicate that frame-related

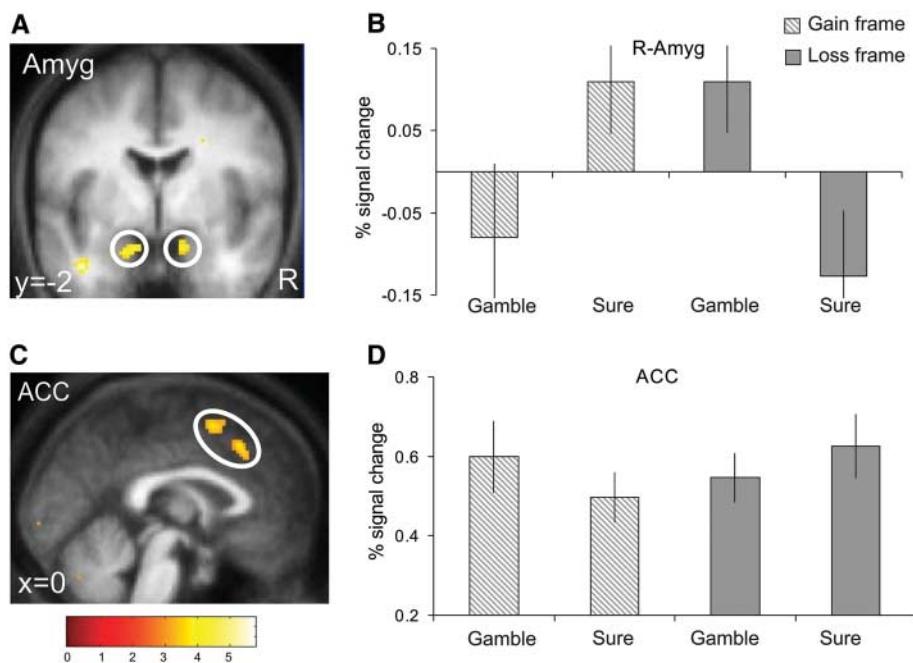
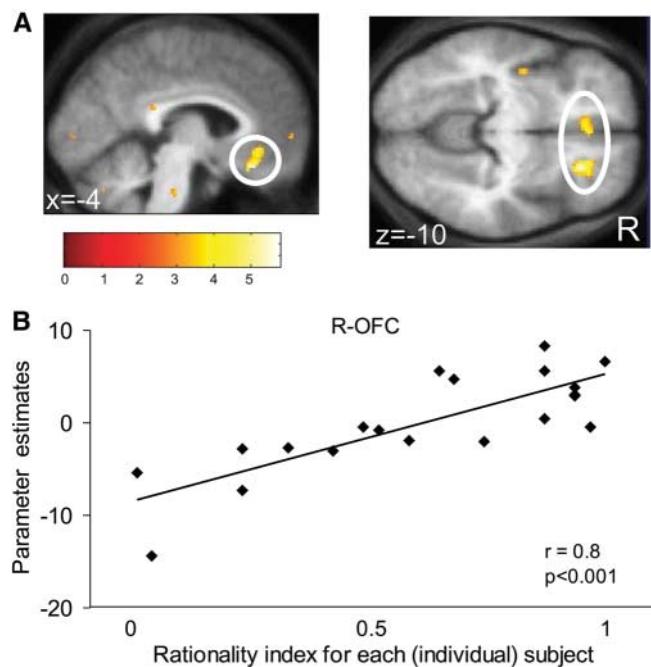


Fig. 3. fMRI results. (A) Interaction contrast [$(G_{\text{sure}} + L_{\text{gamble}}) - (G_{\text{gamble}} + L_{\text{sure}})$]: brain activations reflecting subjects' behavioral tendency to choose the sure option in the Gain frame and the gamble option in the Loss frame (i.e., in accordance with the frame effect). Bilateral amygdala (Amyg) activation [MNI space coordinates (x, y, z): left hemisphere, –14, 2, –24 (peak Z score = 3.97); right hemisphere, 12, 2, –20 (Z score = 3.82)]. (C) Reverse interaction contrast [$(G_{\text{gamble}} + L_{\text{sure}}) - (G_{\text{sure}} + L_{\text{gamble}})$]: brain activations reflecting the decision to choose counter to subjects' general behavioral tendency. Anterior cingulate cortex (ACC) activation: 2, 24, 44 (Z score = 3.65); –2, 8, 56 (Z score = 3.78). Effects in (A) and (C) were significant at $P < 0.001$; for display purposes they are shown at $P < 0.005$. (B) and (D) Plots of percentage signal change for peaks in right amygdala (12, 2, –20) (B) and ACC (2, 24, 44) (D). Error bars denote SEM.

Fig. 4. Rationality across subjects: fMRI correlational analysis. Regions showing a significant correlation between rationality index [between-subjects measure of susceptibility to the framing manipulation; see (14)] and the interaction contrast image [$(G_{\text{sure}} + L_{\text{gamble}}) - (G_{\text{gamble}} + L_{\text{sure}})$] are highlighted. (A) Orbital and medial prefrontal cortex (OMPFC) [MNI space coordinates (x, y, z): VMPFC (left panel), –4, 34, –8 (Z score = 4.56); OMPFC and R-OFC circled in right panel [R-OFC: 24, 30, –10 (Z score = 5.77)]. Effects were significant at $P < 0.001$; for display purposes they are shown at $P < 0.005$. (B) Plot of the correlation of parameter estimates for R-OFC with the rationality index for each subject ($r = 0.8$, $P < 0.001$).



valence information is incorporated into the relative assessment of options to exert control over the apparent risk sensitivity of individual decisions. The observation that the frame has such a pervasive impact on complex decision-making supports an emerging role for the amygdala in decision-making (21, 22).

When subjects' choices ran counter to their general behavioral tendency, there was enhanced activity in the ACC. This suggests an opponency between two neural systems, with ACC activation consistent with the detection of conflict between predominantly "analytic" response tendencies and a more "emotional" amygdala-based system (23, 24).

Previous descriptions of the frame effect have been predominantly confined to between-subjects investigations. Our experimental design allowed us to distinguish the anatomical bases of the frame effect, both within and between subjects. Interestingly, amygdala activity did not predict the substantial intersubject difference in terms of susceptibility to the frame effect. Instead, subjects' tendency to be susceptible to the frame showed a robust correlation with neural activity in the OMPFC. It is noteworthy that there are strong reciprocal connections between the amygdala and the OMPFC (25), although each may contribute to distinct functional roles in decision-making (26). Lesions of the OMPFC cause impairments in decision-making; these are often characterized as an inability to adapt behavioral strategies according to the consequences of decisions, leading to impulsivity (27, 28). It is thought that the OMPFC, incorporating inputs from the amygdala, represents the motivational value of stimuli (or choices), which allows it to integrate and evaluate the incentive value of predicted outcomes in order to guide future behavior (29, 30). Our data raise an intriguing possibility that more "rational" individuals have a better and more refined representation of their own emotional biases that enables them to modify their be-

havior in appropriate circumstances, as for example when such biases might lead to suboptimal decisions. As such, our findings support a model in which the OMPFC evaluates and integrates emotional and cognitive information, thus underpinning more "rational" (i.e., description-invariant) behavior.

Our findings suggest a model in which the framing bias reflects an affect heuristic by which individuals incorporate a potentially broad range of additional emotional information into the decision process. In evolutionary terms, this mechanism may confer a strong advantage, because such contextual cues may carry useful, if not critical, information. Neglecting such information may ignore the subtle social cues that communicate elements of (possibly unconscious) knowledge that allow optimal decisions to be made in a variety of environments. However, in modern society, which contains many symbolic artifacts and where optimal decision-making often requires skills of abstraction and decontextualization, such mechanisms may render human choices irrational (31).

References and Notes

1. K. J. Arrow, *Econ. Inq.* **20**, 1 (1982).
2. A. Tversky, D. Kahneman, *J. Bus.* **59**, S251 (1986).
3. J. von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton Univ. Press, Princeton, NJ, 1944).
4. D. Kahneman, A. Tversky, *Choices, Values, and Frames* (Cambridge Univ. Press, New York, 2000).
5. B. J. McNeil, S. G. Pauker, H. C. Sox Jr., A. Tversky, *N. Engl. J. Med.* **306**, 1259 (1982).
6. D. Kahneman, A. Tversky, *Econometrica* **47**, 263 (1979).
7. A. Tversky, D. Kahneman, *Science* **211**, 453 (1981).
8. H. Damasio, T. Grabowski, R. Frank, A. M. Galaburda, A. R. Damasio, *Science* **264**, 1102 (1994).
9. J. Greene, J. Haidt, *Trends Cogn. Sci.* **6**, 517 (2002).
10. G. F. Loewenstein, E. U. Weber, C. K. Hsee, N. Welch, *Psychol. Bull.* **127**, 267 (2001).
11. T. Gilovich, D. W. Griffin, D. Kahneman, Eds., *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge Univ. Press, New York, 2002).
12. P. Slovic, M. Finucane, E. Perers, D. MacGregor, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, T. Gilovich, D. W. Griffin, D. Kahneman, Eds. (Cambridge Univ. Press, New York, 2002), pp. 397–421.
13. X. Gabaix, D. Laibson, in *The Psychology of Economic Decisions, Vol. 1: Rationality and Well-Being*, I. Brocas, J. D. Carrillo, Eds. (Oxford Univ. Press, Oxford, 2003), pp. 169–183.
14. See supporting material on *Science Online*.
15. M. G. Baxter, E. A. Murray, *Nat. Rev. Neurosci.* **3**, 563 (2002).
16. J. E. LeDoux, *The Emotional Brain* (Simon & Schuster, New York, 1996).
17. J. J. Paton, M. A. Belova, S. E. Morrison, C. D. Salzman, *Nature* **439**, 865 (2006).
18. L. H. Corbit, B. W. Balleine, *J. Neurosci.* **25**, 962 (2005).
19. R. Adolphs, *Brain Res.* **1079**, 25 (2006).
20. H. Kim et al., *J. Cogn. Neurosci.* **16**, 1730 (2004).
21. M. Hsu, M. Bhatt, R. Adolphs, D. Tranel, C. F. Camerer, *Science* **310**, 1680 (2005).
22. B. W. Balleine, S. Kilcross, *Trends Neurosci.* **29**, 272 (2006).
23. M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, J. D. Cohen, *Psychol. Rev.* **108**, 624 (2001).
24. E. K. Miller, J. D. Cohen, *Annu. Rev. Neurosci.* **24**, 167 (2001).
25. D. G. Amaral, J. L. Price, A. Pitkänen, S. T. Carmichael, in *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*, J. P. Aggleton, Ed. (Wiley-Liss, New York, 1992), pp. 1–66.
26. C. A. Winstanley, D. E. Theobald, R. N. Cardinal, T. W. Robbins, *J. Neurosci.* **24**, 4718 (2004).
27. A. Bechara, A. R. Damasio, H. Damasio, S. W. Anderson, *Cognition* **50**, 7 (1994).
28. E. T. Rolls, J. Hornak, D. Wade, J. McGrath, *J. Neurol. Neurosurg. Psychiatry* **57**, 1518 (1994).
29. G. Schoenbaum, B. Setlow, M. P. Saddoris, M. Gallagher, *Neuron* **39**, 855 (2003).
30. G. Schoenbaum, M. R. Roesch, T. A. Stalnaker, *Trends Neurosci.* **29**, 116 (2006).
31. K. Stanovich, F. West, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, T. Gilovich, D. W. Griffin, D. Kahneman, Eds. (Cambridge Univ. Press, New York, 2002), pp. 421–440.
32. Supported by a Wellcome Trust Programme Grant (R.J.D.) and a Wellcome Trust studentship (B.D.M.). We thank H. Spiers, P. Sterzer, and J. Hughes for helpful discussions during the analysis of the study and P. Bossaerts for useful comments on the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/313/5787/684/DC1
Materials and Methods

Figs. S1 to S3

Tables S1 to S3

References

5 April 2006; accepted 12 June 2006
10.1126/science.1128356

Anterior cingulate reflects susceptibility to framing during attractiveness evaluation

Michael Deppe^{a,*}, Wolfram Schwindt^{b,*}, Anna Pieper^a, Harald Kugel^b, Hilke Plassmann^c, Peter Kenning^c, Katja Deppe^d and E. Bernd Ringelstein^a

Departments of ^aNeurology, ^bClinical Radiology, ^cGeneral Management, University of Münster and University Hospital Münster (UKM) and ^dDepartment of Neurology, Franz Hospital Dülmen, Germany

Correspondence to Dr Michael Deppe, PhD, Department of Neurology, University of Münster, Albert-Schweitzer-Str. 33, D-48129 Münster, Germany
Tel: +49 (0) 251/83 48174; fax: +49 (0) 251/83-52064; e-mail: mail@Michael-Deppe.de

*Michael Deppe and Wolfram Schwindt contributed equally to this study.

Received 19 April 2007; accepted 22 April 2007

Human cognitive decisions can be strongly susceptible to the manner in which options are presented ('framing effect'). Here we investigated the neural basis of response adjustments induced by changing frames during intuitive decisions. Evidence exists that the anterior cingulate cortex plays a general role in behavioral adjustments. We hypothesized, therefore, that the anterior cingulate cortex is also involved in the 'framing effect'. Our hypothesis was tested by using a binary attractiveness judgment task ('liking'

versus 'nonliking') during functional magnetic resonance imaging. We found that the framing-related anterior cingulate cortex activity predicted how strongly susceptible an individual was to a biased response. Our results support the hypothesis that paralimbic processes are crucial for predicting an individual's susceptibility to framing. *NeuroReport* 18:II19–II23 © 2007 Lippincott Williams & Wilkins.

Keywords: decision-making, fMRI, framing effect, neuroeconomics, prefrontal cortex, prejudgments, uncertainty

Introduction

In decision science, the term 'framing' characterizes the manner in which a problem of choice is presented. For example, the judgment of whether a certain newsmagazine headline is regarded as true or as false strongly depends on the credibility of the magazine in which the headline is presented [1]. In this example, the magazine brand can be regarded as 'framing information' that frames a cognitive plausibility decision. The dependence of the answers during a choice task on framing has been denoted as 'framing effect'. Tversky and Kahneman's work on the perceptual principles of human cognitive-decision problems has demonstrated that economic decisions can strongly depend on framing [2].

Recent studies provide evidence that framing-induced judgment biases are associated with emotional information integrated into the cognitive-decision process [1,3]. In particular, it was shown that during rational decision-making activations in the prefrontal cortex correlated with the participants' susceptibilities to framing effects [1].

From a phylogenetic point of view, the development of adaptive behavior requires that all kinds of environmental stimuli be evaluated with respect to their behavioral relevance. Thus, framing effects might occur not only during rational decision-making with a purely cognitive context, but also during intuitive behavior, even if no obvious relationship exists between the framing information and the judgment material.

The anterior cingulate cortex is a central node in a neural network responsible for the integration of information about positive or negative reinforcements. This node thus relates actions to consequences, and is central to conflict monitoring and cognitive control [4–6]. Evidence from human and animal studies is available about the role of the anterior cingulate cortex in assessing the need for behavioral adaptations [7–10].

In this study we tested the hypothesis that the anterior cingulate cortex also has a central function for the integration of framing information into response behavior during intuitive decision making. When being investigated by functional magnetic resonance imaging (fMRI), participants had to evaluate the attractiveness ('liking', 'nonliking') of print advertising. The framing information consisted of the logos and style elements of four German newsmagazines and was concomitantly presented with the advertisements during the judgment task.

Methods

Stimulus material

We employed 46 current print advertisements taken arbitrarily from a well known German newsmagazine (Focus Magazin Verlag, Hubert Burda Media GmbH, München, Germany). In a pilot study, 100 randomly selected participants of both sexes had to judge their liking of these advertisements. The advertisements had to be rated on a

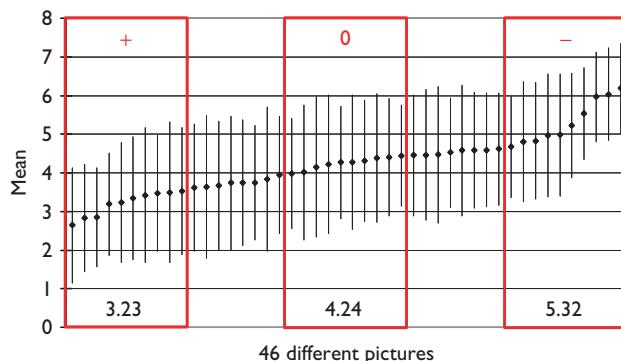


Fig. 1 Mean ratings and standard deviations of 46 existing newsmagazine ads from 100 participants assessed by questionnaires. The ratings were graded into seven categories (1 = very attractive, to 7 = absolutely unattractive). The 10 most attractive (group '+'), the 10 most nonliking (group '-'), and 10 indifferent ads (group '0') were selected for our experiment. By using this kind of 'normalization,' it was expected that the statements of the three groups would mainly be judged as 'liking', 'ambiguous' and 'nonliking', respectively.

scale ranging from 1 = 'very attractive' to 7 = 'absolutely not attractive' (see Fig. 1). As stimulus material for the following fMRI experiments we selected 30 of these advertisements, namely 10 advertisements each in the groups rated highest (group '+'), lowest (group '-'), or indifferent (group '0'), respectively (see red boxes, Fig. 1).

Participants undergoing fMRI

Thirteen healthy men and eight healthy women (these participants did not belong to the pilot study) participated in the core trial (mean age 29, range 25–35 years). Standard exclusion criteria for MR imaging were applied. Participants with strong myopia or other relevant constraints of vision were also excluded. All participants provided written informed consent before the scanning sessions. The participants were also informed that the examination could reveal potentially medically significant findings and were asked if they would like to be notified in this case. The study was approved by the local ethics committee.

Judgment task

During fMRI the 30 advertisements were presented in combination with each of four logos of well known German newsmagazines, here denoted as A, B, C, and D (Fig. 2). Thus 4×30 judgments had to be made. For data analysis, these logos were used as the four levels of the framing factor. The interstimulus interval between all 120 stimuli was standardized to 10 s. In a forced-choice task, the participants were requested to judge whether they liked the presented advertisement or not by pressing the corresponding button ('liking' or 'nonliking') on a response box compatible with the MR scanner. The responses were recorded by the stimulation software that was used for visual presentation (ShowPics, Department of Neurology, University of Münster, Germany). Separately for each magazine M , we calculated the response bias B_M according to [1]

$$B_M = \frac{N_{\text{Liking}} - N_{\text{Nonliking}}}{N_{\text{Liking}} + N_{\text{Nonliking}}}, \quad (1)$$

where N_{Liking} and $N_{\text{Nonliking}}$ are the number of ratings judged as 'liking' or 'nonliking', respectively ($-1 < B_M < 1$). Owing to



Fig. 2 Visual presentation of the Attractiveness (3) \times Frame (4) within-participant factorial design for the judgment task (forced answers 'liking', 'nonliking'). In addition to the advertising, which represented the explicit decision-relevant information, one of the four different logos A, B, C, or D and the corresponding style elements of four different German newsmagazines were presented as framing stimuli.

the preceding normalization of the stimulus material, we assumed an average bias B of around 0 for representative participants, whereas 'ad-liking' participants would obtain positive and 'ad-aversive' persons negative values for B .

Additionally, a susceptibility index SI was calculated as a measure for the variability of the responses with regard to the randomly changed magazine logos $M = A, B, C$, and D :

$$SI = \sqrt{\frac{1}{4} \sum_{M=A,B,C,D} (\bar{B} - B_M)^2}. \quad (2)$$

Mathematically the SI is the standard deviation of the response biases B_A, B_B, B_C and B_D for each individual using the mean bias

$$\bar{B} = (B_A + B_B + B_C + B_D)/4. \quad (3)$$

Accordingly, if the magazine brand had no influence on the perception and processing of the advertisements, there would be no differences between the four B_M values, that is, $B_A = B_B = B_C = B_D$, and thus $SI = 0$. SI can range between 0 and 1. $SI = 1$ indicates that the response variability is entirely due to framing manipulation, that is, the visual content of all advertisements is completely ignored.

Image presentation

A dedicated fMRI projection system (Covilex, Magdeburg, Germany) provided high-quality image presentation with the images covering about 50% of the participant's whole field of view. Even small details of the visual stimulus could be recognized easily. Care was taken to present the different advertisements and logos in equal size, position, background, and luminance to prevent such potential confounders. The volunteers were asked to press the corresponding button on the response box immediately after they had made their decisions. They were informed that 120 decisions had to be made. Head fixation was achieved with foam pads and a soft headband. Earplugs and headsets were employed to protect against scanner noise and to permit communication with the participants.

Magnetic resonance image acquisition

All data were acquired on a 3.0-T whole-body scanner (Intera T30, Philips, Best, The Netherlands). An isotropic T1w dataset of the whole head with a reconstructed voxel size of 0.5-mm edge length was acquired for anatomic identification and coregistration into the Talairach space [11]. For functional images blood oxygenation level dependent (BOLD) contrast images were acquired using a T2*-weighted single shot gradient echo-planar imaging (EPI) sequence, which covered nearly the whole brain. The dataset consisted of 36 transversal slices of 3.6-mm thickness without gap, field of view 230×230 , matrix 64×64 . Slices were oriented parallel to the ac-pc line. Imaging parameters were repetition time = 3000 ms, echo time = 50 ms, flip angle = 90°, and EPI-factor (echo train length) 63. Before each fMRI was run, 10 dummy scans were acquired to allow for the equilibration of magnetization.

Data analysis

Data analysis was performed using Statistical Parametric Mapping (SPM2; Wellcome Department of Cognitive Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>). All EPI volumes were spatially normalized to the Montreal Neurologic Institute EPI standard template and resampled to $2 \times 2 \times 2$ mm³ resolution [12]. All normalized functional volumes were smoothed with an isotropic Gaussian kernel (4 mm full width at half-maximum). Global changes in fMRI response from scan to scan were removed by proportional scaling to have a common global mean voxel value. To correct for long-term effects, we carried out high-pass filtering with a cutoff frequency of 0.008 Hz. The hemodynamic responses were modeled into an event-related statistical design based on the general linear model. The temporal events were defined by the time points when the visual stimuli appeared. The events were categorized according to the four different levels of the implicit stimulation, that is, the occurrence of the advertisements in one of the four different newsmagazine brands A, B, C, and D. According to these four levels of the factor magazine type *M*, a one-factorial analysis of variance (ANOVA) was calculated for each examined participant to assess individual cortical-activity modulations caused by the framing factor (single-participant analysis). For the group analysis, we employed linear regression by using the *F* maps from the one-factorial single-participant ANOVA as the dependent variable and the participants' *SI* as the independent variable. Correlations of $P < 0.001$ (uncorrected, extent threshold = 50 voxels) between *SI* and cortical-activity modulations were regarded as being significant.

Results

Judgments

Owing to the 'normalization' of the advertisements' attractiveness (Fig. 1) in the pilot study, we found balanced decisions in total. The actual judgments averaged over the four journals were mainly 'liking' for the '+' group, 'nonliking' for the '-' group, and were nearly equally distributed within the '0' group. According to equation 1, this led to an overall mean bias for all 21 examined participants close to zero. Response times typically varied between 2 and 4 s with significantly longer times for decisions of the '0' group. Our behavioral results expressed by the *SI* [equation (2)] indicate a violation of representa-

tional invariance in response behavior due to the framing by the magazine logos. *SI* varied between 0.0 and 0.32, indicating an interindividual difference in the degree of susceptibility for framing information (Fig. 3).

Cortical activity

The fMRI analysis performed at the single-participant level (one-factorial ANOVA) revealed significant blood oxygenation level dependent signal modulations by framing for about half the participants. In participants with significant effects, these changes occurred consistently in the right-sided higher visual areas in the parietal cortex [Brodmann area (BA) 19], in the inferior frontal gyrus (BA 46), the right middle frontal gyrus (BA 10), and bilaterally within the anterior cingulate cortex, in correspondence with our hypothesis (Fig. 4a-c).

Regression analysis demonstrated that the degree of the participants' susceptibilities to judgment biases, expressed by the *SI*, and their corresponding brain-activity modulations in the anterior cingulate cortex, induced by the varying magazine brands, correlated highly significantly ($R = 0.69$, $P = 0.0005$) (Fig. 4d).

Discussion

In two recent studies employing (i) a preference-choice task [13] and (ii) a cognitive plausibility-evaluation task [1], it was shown that the ventral parts of the medial prefrontal cortex play a key role in the integration of implicit, decision-relevant framing information during rational decision making. These structures have been found to be particularly involved in processing emotions. Further, it has been demonstrated that framing effects correlated with activity modulations in emotionally associated brain structures during economical decisions [3].

In this study, we used the same framing stimulus as in the previously published plausibility-evaluation task [1], but we

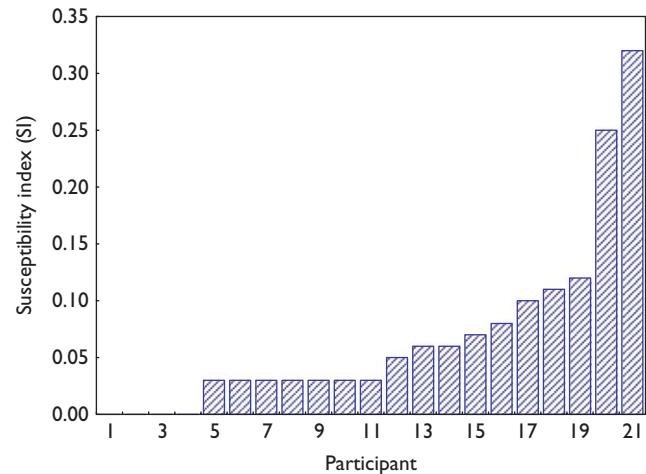


Fig. 3 Behavioral results. Each bar represents the susceptibility index (*SI*) of each individual participant to the framing information presented during the judgments. *SI* was computed from the participant's responses and reflects the individual's susceptibility to framing. *SI* can hypothetically range between 0 and 1; *SI* = 0 would imply absolutely no influence of framing on responses. In other words, identical advertisements were judged equally in all four frames, *SI* = 1 would indicate that response variability is entirely due to framing manipulation, that is, the visual content of all advertisements was completely ignored.

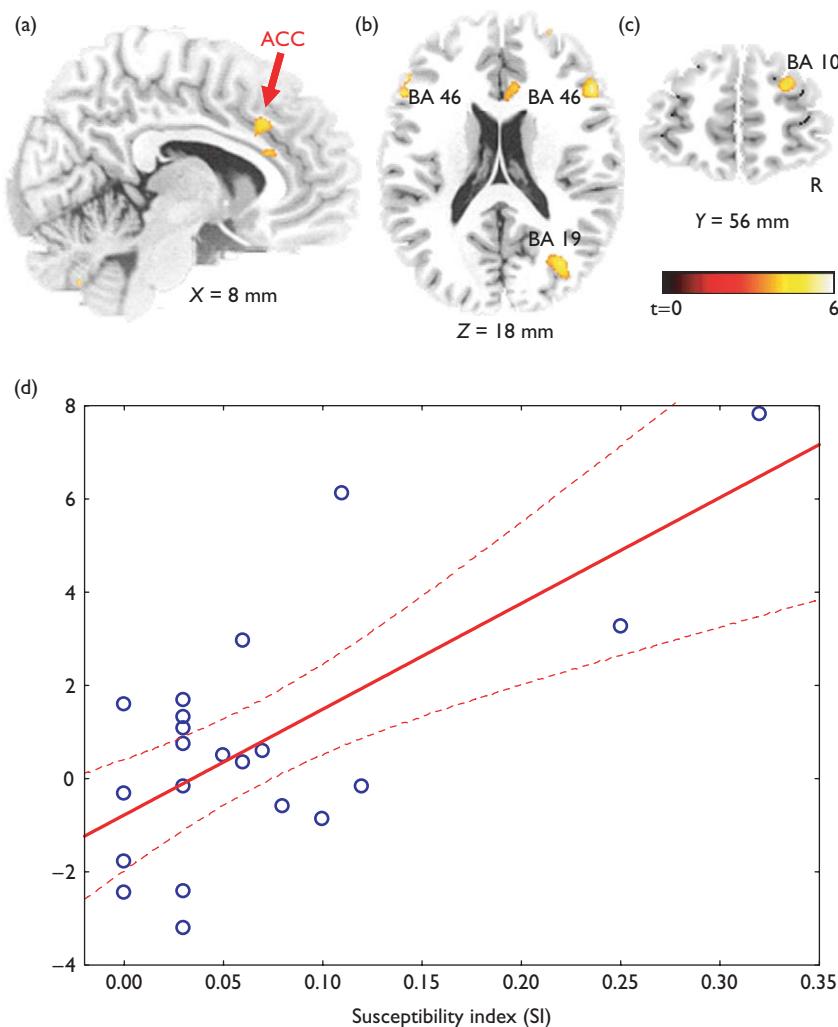


Fig. 4 Stimulus-onset-locked brain-activity changes. (a–c) We found significant correlations ($P < 0.001$, uncorrected, extent threshold 50 voxels) between individuals' susceptibility for a framing effect (SI) and the effect size of the framing stimulus on the hemodynamic response bilaterally in the anterior cingulate cortex (red arrow, a) and the inferior frontal gyrus, Brodmann area (BA) 46 (b). Further correlations were found in higher-order visual cortices in the right occipital lobe (BA 19), and the right middle frontal gyrus, BA 10 (c). (d) Linear regression between the response effect in the anterior cingulate cortex (ACC; red arrow in a) and SI. Each point reflects data from a single participant ($N = 21$).

changed the type of material that had to be evaluated. Instead of semantic stimuli, we used visual stimuli here, which had to be intuitively evaluated on a participant-specific intrinsic scale. We further assessed the event-related hemodynamic responses immediately associated with the occurrence of the visual stimulus instead of those at the point in time when the responses were given.

The main finding of this study was that anterior cingulate cortex activity predicted the individual's susceptibility to framing information. If the magazine brand had no influence on the perception and processing of the advertisements in the sense of a framing effect, there would have been no significant differences between the four B_M values, that is $B_A = B_B = B_C = B_D$, and thus $SI = 0$ (according to equation 2), because the visual contents to be evaluated were exactly the same in all four magazines. Quite the contrary was true. We found highly significant intraindividual variations in B_M (expressed by SI), reflecting a strong influence of the confounding framing stimuli, that is, the magazine brands. The strong correlation between an

individual's SI score, representing a participant's susceptibility to framing, and the activity modulation in the anterior cingulate cortex supports our hypothesis that this paralimbic structure is involved in assessing the behavioral relevance of framing stimuli. These results are in agreement with recent findings that underline the role of the dorsal anterior cingulate cortex in stimulus assessment and in the general regulation of response behavior beyond conflict monitoring [14].

Conclusion

We conclude that, corresponding to its fundamental role in behavioral adjustments, the anterior cingulate cortex is involved in the assessment of environmental (framing) information during decision making. Our results support the hypothesis that an individual's susceptibility to framing is at least partly regulated by presemantic assessment processes in a phylogenetically old brain structure. These results may explain why it is so hard to fulfill an important

axiom in rational decision-making theory, namely the representational invariance in human decision behavior [15].

Acknowledgements

The authors acknowledge the invaluable help of Benjamin Gess and Dipl.-Phys. Jens Sommer in supporting the development of the stimulation design. This work was supported by grants of the Stiftung Neuromedizin Medical Foundation. Stimulus material and financial support was provided by Hubert Burda Media GmbH, München, Germany (Dr Anna Maria Deisenberg, Michael Pusler).

References

1. Deppe M, Schwindt W, Kramer J, Kugel H, Plassmann H, Kenning P, et al. Evidence for a neural correlate of a framing effect: bias-specific activity in the ventromedial prefrontal cortex during credibility judgments. *Brain Res Bull* 2005; **67**:413–421.
2. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science* 1981; **211**:453–458.
3. De Martino B, Kumaran D, Seymour B, Dolan RJ. Frames, biases, and rational decision-making in the human brain. *Science* 2006; **313**:684–687.
4. Egner T, Hirsch J. The neural correlates and functional integration of cognitive control in a Stroop task. *Neuroimage* 2005; **24**:539–547.
5. Kerns JG, Cohen JD, MacDonald AW III, Cho RY, Stenger VA, Carter CS. Anterior cingulate conflict monitoring and adjustments in control. *Science* 2004; **303**:1023–1026.
6. Rushworth MF, Walton ME, Kennerley SW, Bannerman DM. Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci* 2004; **8**:410–417.
7. Eisenberger NI, Lieberman MD, Williams KD. Does rejection hurt? An fMRI study of social exclusion. *Science* 2003; **302**:290–292.
8. Walton ME, Bannerman DM, Alterescu K, Rushworth MF. Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J Neurosci* 2003; **23**:6475–6479.
9. Walton ME, Bannerman DM, Rushworth MF. The role of rat medial frontal cortex in effort-based decision making. *J Neurosci* 2002; **22**:10996–11003.
10. Walton ME, Devlin JT, Rushworth MF. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat Neurosci* 2004; **7**:1259–1265.
11. Deppe M, Steinrätter O, Knecht S. The T2T-database Java applet. *Neuroimage* 2003; **19**:S48.
12. Ashburner J, Neelin P, Collins DL, Evans A, Friston K. Incorporating prior knowledge into image registration. *Neuroimage* 1997; **6**:344–352.
13. Deppe M, Schwindt W, Kugel H, Plassmann H, Kenning P. Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J Neuroimaging* 2005; **15**:171–182.
14. Roelofs A, van Turennout M, Coles MG. Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proc Natl Acad Sci U S A* 2006; **103**:13884–13889.
15. Von Neumann J, Morgenstern O. *Theory of games and economic behavior*. Princeton NJ: Princeton University Press; 1944.

How choice ambiguity modulates activity in brain areas representing brand preference: evidence from consumer neuroscience

Hilke Plassmann^{1*}, Peter Kenning², Michael Deppe³,
Harald Kugel⁴ and Wolfram Schwindt⁴

¹Division of Humanities and Social Sciences, California Institute of Technology, USA

²Department of Marketing, Zeppelin University, Germany

³Department of Neurology, University of Muenster and University Hospital Muenster, Germany

⁴Department of Clinical Radiology, University of Muenster and University Hospital Muenster, Germany

- In the present paper we investigate whether choice ambiguity modulates activity in brain areas that represent brand preference and decision utility, as identified in previous studies. Our findings reveal that brain areas involved in the interaction of brand information and ambiguity information are the (predominantly left) ventromedial prefrontal cortex (VMPFC) and the anterior cingulate (AC). These activation patterns have earlier been found to correlate with brand preference. Thus, our findings show that the reduction of perceived ambiguity and information costs by brand information drives neural representations of brand preference as promoted by signaling theory in information economics.

Copyright © 2008 John Wiley & Sons, Ltd.

Introduction

Previous research in marketing provides evidence that knowledge of the brand's name is an important product attribute, influencing consumers product perceptions and, through these their product preferences, their willingness to pay, consumption experiences, and pro-

duct loyalty (Allison and Uhl, 1964; Aaker, 1992; Keller, 1993; de Chernatony and McDonald, 1998). For example, a study from Allison and Uhl (1964) shows that consumers like the taste of a beer significantly more when they know the name of the beer brand, as compared to an anonymous tasting. The influence of brand name information on consumer choice seems to be particularly important in the case of product choices that are very similar with respect to their core attributes such as quality, price, and utilitarian value proposition. Other studies suggest that this is due to the signaling function of the brand name that triggers a variety of different

*Correspondence to: Hilke Plassmann, Division of Humanities and Social Sciences, California Institute of Technology, USA.

E-mail: hilke_plassmann@hss.caltech.edu

Note: All figures will be produced in colour in the online published version.

associations like prior experiences with the brand, brand-quality inferences, and that reduces information asymmetry and information/search costs (Erdem and Swait, 1998).

Even more intriguingly, research in the nascent field of consumer neuroscience (Shiv *et al.*, 2005; Kenning and Plassmann, 2008; Lee *et al.*, 2007) indicates that brand preferences recruit specific brain systems during both choice tasks and the consumption experience. For example, a functional magnetic resonance imaging (fMRI) study by Deppe *et al.* (2005) showed that decision-making with prior brand preferences led to increased activity in brain regions involved in emotional and self-referential processing such as the ventromedial part of the prefrontal cortex (VMPFC) and the anterior cingulate cortex (ACC). In contrast, choices without prior preferences evoked increased activations of brain areas that are also activated during working memory tasks, introspection, and visual processing. Another fMRI study by McClure *et al.* (2004) shows that activity within the VMPFC correlates with prior preferences for colas during anonymous tasting. In addition, they found that the underlying mental process in the case when people know which cola brand they taste, as compared to uncertainty about which cola they taste, recruits memory areas of the brain such as the hippocampus and the dorsolateral prefrontal cortex (DLPFC). The authors suggest that subjects are retrieving their prior experiences with the brand during their consumption experience. In addition, a study involving patients with lesions in the VMPFC suggests that a lesion in the VMPFC, as compared to lesions in other areas of the brain and to healthy subjects, decreases the brand information bias on preference formation (Koenigs and Tranel, 2007). The commonalities of all these studies are the use of nearly identical products as stimuli and brand names as the main discriminative factor.

In the present paper we take these studies a step further by investigating the influence of ambiguity on the signaling power of brand information. Hence, we investigated whether (and, if yes, how) choice ambiguity modulates

activity in brain areas that were found to be involved in processing brand preference and decision utility, in accordance with previous studies [(i.e., the VMPFC and medial orbitofrontal cortex; Paulus and Frank, 2002; Erk *et al.*, 2002; Deppe *et al.*, 2005; Plassmann *et al.*, 2007)]. We scanned the brains of 15 subjects as they were making choices between nearly identical vacation packages (with only semantic differences in the description) to different destinations (low ambiguity: Germany's Black Forest, high ambiguity: Israel) offered by different travel operator brands (e.g., Thomas Cook vs. TUI). We found evidence for an increased preference signaling function of brand information in a high as compared to a low ambiguity context, accompanied by increased changes in neural activation changes in the VMPFC and the ACC.

The remainder of the paper is organized as follows. In the next section we will provide the conceptual background for our study that compares predictions based on signaling theory of information economics with those from branding theories rooted in cognitive psychology. Then we will describe our empirical study and conclude with a discussion of our results.

Conceptual background

Most branding theories are derived from cognitive psychology. They suggest that brand awareness, brand associations, and perceived quality lead to higher customer-based brand equity and, in turn, reduce perceived ambiguity (Aaker, 1992; Chaudhuri and Holbrook, 2001; Keller, 1993, 2002, 2003). These approaches predict that during choice under ambiguity the mere brand information triggers the modulation of neuronal activity in areas shown to represent brand preference such as the VMPFC (McClure *et al.*, 2004; Deppe *et al.*, 2005; Koenigs and Tranel, 2007).

Other theoretical approaches are based on signaling theory from information economics (e.g., Stigler, 1961; Stiglitz, 1989). They suggest that brand information serves as a signal of tangible (e.g., product quality) and intangible

(e.g., hedonic/social) utility that reduces asymmetrical information and then results in brand equity and preference (Erdem and Swait, 1998, 2004; Krishnan and Hartline, 2001). In other words, approaches from information economics propose that reduction of perceived ambiguity and information costs drives brand preference (Erdem and Swait, 1998, 2004). Thus, these approaches would predict that a reduction of asymmetric information due to an interaction of ambiguity and brand information signal would recruit brain areas involved in processing brand preference such as the VMPFC (McClure *et al.*, 2004; Deppe *et al.*, 2005; Koenigs and Tranel, 2007).

Against this background we tested the hypothesis that mere brand information can trigger activation found to represent brand preference during choice under ambiguity, as suggested by brand theories based on cognitive psychology (i.e., a main effect of brand information) against the alternative hypothesis that neural activation correlating with brand preference is triggered, instead, by the brand information and potential to reduce ambiguity (i.e., an interaction effect of brand information and high vs. low ambiguity), as promoted by signaling theory.

Empirical study

In order to investigate this hypothesis, we first conducted two pretests to select the stimulus material for the subsequent main study.

Pretests

In a first behavioral pretest we investigated whether brand information is an important service attribute when customers book a vacation package from a certain travel operator. In order to keep product attributes other than the brand information constant, the hypothetical offers consisted of nearly identical travel services that differed only in the wording of their description. In a two-staged experiment similar to anonymous taste tests of Coke versus Pepsi (de Chernatony and McDonald, 1998) we were able to show that as

compared to anonymous tasting (stage 1), the knowledge of brand information (stage 2) led to preference reversals (see **Figure 1 A**, $N=25$).

We conclude that the factor "brand information" is important for travel service decisions and can be separated mentally from the service itself.

In a second pretest we tested whether ambiguity due to higher credence qualities of goods plays a role for travel services. As a measure for ambiguity we sampled ratings to which extent brand information can signal trust in a product category with low credence qualities (beer and coffee, fast moving consumer goods = FMCG) and compared it with brand trust ratings for travel services. Brand trust was assessed by a direct 5-point measurement scale (1 = very low trust, 5 = very high trust, $N=35$). We employed a one-way analysis of variance (ANOVA) to compare brand trust ratings between these two product categories. We revealed that brand trust in the product category with low credence qualities is significantly lower ($M_{BTrust_FMCG} = 1.8$) than for travel services ($M_{BTrust_Travel_Services} = 4.0$, $F(1, 34) = 10.5$, $p < .005$). These findings support the position that ambiguity due to higher credence qualities plays a role for travel services.

Main study

Task and design

To investigate the impact of brand information signals on choices under ambiguity we scanned human subject brains ($N=15$) using fMRI while the subjects made binary choices between 16 different travel operator brands. Subjects were instructed to choose the travel operator brand from which they would book a travel service to one of the two different destinations. For example, they could choose to book a vacation package to Germany's Black Forest with TUI or Thomas Cook. The design enabled us to compare changes in brain activity in one experimental condition (e.g., no preferred brand for choice) as compared to

another experimental condition (e.g., preferred brand for choice) (the so-called subtraction method in neuroscience). No “resting” condition was implemented as a control condition to avoid the problem of ambiguous baseline conditions (Stark and Squire, 2001). This is in line with other similar studies in consumer neuroscience (e.g., Deppe *et al.*, 2005a, Deppe *et al.*, 2005b). In order to assess reproducibility and habituation effects, one session of 64 choice trials was presented four times to each participant (Lohmann *et al.*, 2004). Further, to avoid response-related brain activations and movement artifacts (Deppe *et al.*, 2005), the respondents were not asked for feedback during the fMRI data acquisition.

As a variation of Deppe *et al.*'s (2005) brand choice task, we added a second factor. We employed a two-factorial blocked design (factor 1: brand, levels = favorite (F), diverse (D); factor 2: ambiguity, levels = high (H), low (L)) resulting in four types of decisions in a pseudo-randomized order: (FD-H): the favorite brand versus various diverse brands and high-ambiguity context; (FD-L): the favorite brand versus various diverse brands and low-ambiguity context; (DD-H): diverse versus other diverse brands and high-ambiguity context; and (DD-L): diverse versus other diverse brands and low ambiguity context (see Figure 1B).

Subjects and stimuli

Fifteen healthy and right-handed subjects participated in the experiment (seven males, eight females, median age 24). All subjects met the requirements to participate in an fMRI study. Subjects were informed about the experiment and gave written consent before participating.

Based on the results of the pretests we selected stimuli from the travel service sector. We used 16 familiar travel operator brands to model the factor brand, and two different travel destinations that are generally perceived as either dangerous (high ambiguity) or not dangerous (low ambiguity) to model the factor ambiguity.

Data acquisition and preprocessing

All data were acquired from a 3.0 Tesla whole body scanner (Intera T30, Philips, Best, NL). For the acquisition of functional images blood oxygenation level dependent (BOLD) contrast images were acquired using a T2* weighted single shot gradient echo-planar imaging (EPI) sequence which covered nearly the whole brain. The data set consisted of 36 transversal slices of 3.6 mm thickness without gap that could be acquired with a TR of 3 seconds. The total acquisition time was 3 minute 20 seconds for each of the four runs. A T1-weighted structural image was also acquired for each subject. In combination with the fMRI measurement, while outside the scanner the participants were asked to rank the relevant travel operator brands according to their preferences outside the scanner.

The fMRI data analysis was performed using the Statistical Parametric Mapping software (SPM2; Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm/). We applied the following preprocessing steps to the imaging data: (1) slice-timing correction (centered at TR/2); (2) realignment to the middle volume; (3) spatial normalization to a standard T2* template with a resampled voxel size of 3.6 mm³; (4) spatial smoothing using a Gaussian kernel with full width at half maximum of 8 mm; and (5) intensity normalization and high pass temporal filtering (filter width 128 s). The structural T1 images were co-registered to the mean functional EPI images for each subject and normalized using parameters derived from the EPI images.

Statistical analysis and results

We estimated a general linear model in which the exposure to each experimental condition was entered as a regressor of interest (Frackowiak *et al.*, 2004). These regressors, plus additional regressors of no interest were convolved with a canonical hemodynamic response function (HRF). We then calculated first-level single-subject contrasts to compare

Table 1. Results for the contrast [FD-H – FD-L] – [DD-H – DD-L], $p = 0.001$ (corr.), $k = 5$ voxels

Cortical region	Side	BA*	t-value	MNI coordinates (x, y, z)
Frontal lobe, middle frontal gyrus	L	10	4.96	-14, 54, 0
Limbic lobe, anterior cingulate	R/L	24	n.s.	0, 36, 4
Limbic lobe, anterior cingulate	L	32	3.00	-14, 32, 25
Parietal/occipital lobe, precuneus	R	7/31	n. s.	14, -61, 29
Frontal lobe, superior frontal gyrus	R	10	n. s.	7, 58, -7

the choice between each individual subject's favorite brand (F) and a non-favorite brand (D) by ambiguity level (H = high, L = low). For these first-level contrasts we calculated two second-level group contrasts using a random-

effects model. This statistical approach of multi-level mixed-effects analysis is widely used in recent fMRI studies and it enables to make the intra- and inter-subject variability considerations (Frackowiak *et al.* 2004).

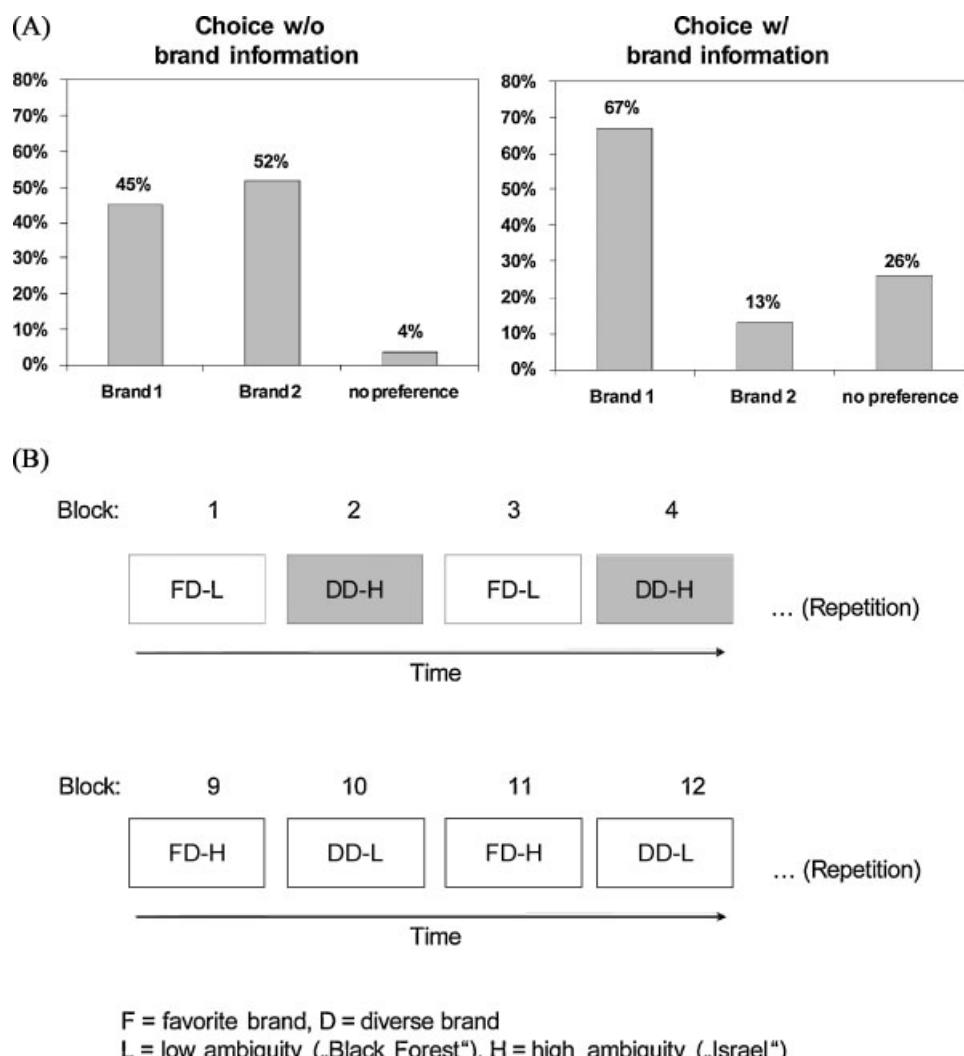


Figure 1. (A) Results of pretest choice of vacation package without and with brand information. (B) Experimental design main study.

First, we employed *t*-tests in a random effects model to analyze the main effect of the first factor (brand information, contrast [FD-H + FD-L] - [DD-H + DD-L]). The statistical data analysis did not result in any significant empirical evidence that the mere brand information triggered activation in brain areas representing brand preference signal.

Second, we employed *t*-tests in a random effects model to test for interaction effects between the factor brand information and the factor ambiguity (contrast [FD-H - FD-L] - [DD-H - DD-L]). We found significant interactions ($p < 0.0001$, corrected, minimum cluster size 10 voxels) reflected by activity changes within the left and right ventromedial prefrontal cortex (BA 10), the left anterior

cingulate (BA 32), and the right precuneus in the occipital/parietal lobe (BA 31) (see **Table 1** and **Figure 2**).

Discussion

The purpose of this study was to investigate the role of choice ambiguity for the neural representation of brand preference. Based on prior research, it was assumed that specific brand information acts as marker signals that correlate with activity changes in areas encoding brand preference such as the VMPFC (McClure *et al.*, 2004; Deppe *et al.*, 2005; Koenigs and Tranel, 2007).

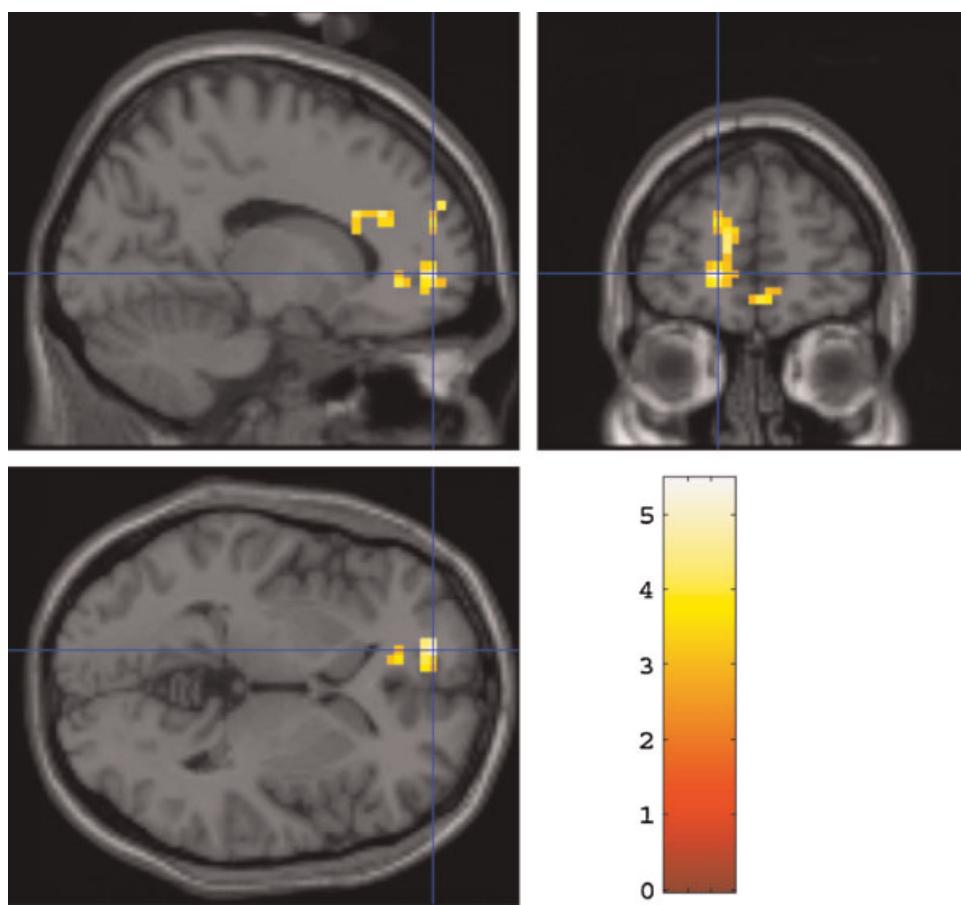


Figure 2. Activity in the VMPFC was higher for condition in the high versus low ambiguity context when the subject's favorite brand was present [FD-H - FD-L] than in the condition in the high versus low ambiguity context when the subject's favorite brand was absent [DD-H - DD-L]. Results for the contrast [FD-H - FD-L] - [DD-H - DD-L]. Activation maps are shown at a threshold of $p < 0.001$ corrected and with an extend threshold of 10 voxels. Scale encodes magnitude of *t*-values.

We found that mere brand information in a more complex choice context does not activate brain areas representing brand preference. This finding contradicts assumptions developed from cognitive psychology brand theories that would predict that, during choice under ambiguity, the mere brand information triggers the modulation of neuronal activity in areas in which activity changes were shown to correlate with brand preference, which in turn reduces perceived ambiguity. The present paper reveals that brain areas involved in the interaction of brand information and ambiguity information are the (predominantly left) ventromedial prefrontal cortex (VMPFC) and anterior cingulate (AC). These activation patterns have been found to correlate with brand preference (McClure *et al.*, 2004; Deppe *et al.*, 2005; Koenigs and Tranel, 2007). Our findings support the hypothesis from signaling theory that the reduction of perceived ambiguity and information costs by brand information drives neural representations of brand preference. Thus, the present findings bridge the gap between the two theoretical approaches in branding theory.

Acknowledgements

We gratefully acknowledge support by Jens Sommer.

Biographical notes

Hilke Plassmann is postdoctoral researcher in the Division of Humanities and Social Sciences of the California Institute of Technology, where she has been on the faculty since July 2006. Before that, she was a postdoctoral researcher in the Stanford Neuroeconomics Lab, Stanford University. She received a Ph.D. (Marketing and Neuroscience) from the University of Muenster's School of Business and Economics in 2005. Her primary research areas are consumer decision-making, self-control, and neuroeconomics. Her work has been published in various marketing and neuroscience journals and has been presented

at diverse conferences focusing on behavioral decision research.

Peter Kenning is Professor of Marketing at Zeppelin University, Germany. His primary research interests are consumer behavior, consumer neuroscience, neuroeconomics, and marketing management. His work has been published widely, e.g., in *Management Decisions*, *Journal of Product and Brand Management*, and *Advances in Consumer Research*, as well as in the *Journal of Neuroimaging*, *NeuroReport*, and *Brain Research Bulletin*. He has presented papers at numerous conferences, including AMA, EMAC, and ACR. For his work he has received several best-paper awards and grants from the German government.

Michael Deppe is physicist (Ph.D.) and senior researcher in the Department of Neurology at the University of Muenster.

Harald Kugel is physicist (Ph.D.) and senior researcher in the Department of Clinical Radiology at the University of Muenster. His fields of research comprise advanced applications of magnetic resonance techniques, especially functional magnetic resonance imaging and magnetic resonance spectroscopy with a focus on normal and pathological brain function. He is author or co-author of more than 100 scientific papers.

Wolfram Schwindt is radiologist and postdoctoral researcher in the Department of Radiology, University Hospital Münster. After studying medicine in Münster he started his internship in internal medicine in Münster University Hospital. He wrote his doctoral thesis in the Institute of Physiology (epilepsy research). After the internship he spent one year as a postdoctoral fellow in the Department of Physiology and Neuroscience, NYU, New York and received further training as a postdoc in the Max-Planck-Institute for Neurological Research, Cologne, where he worked in stroke research and on the physiological mechanisms underlying functional brain activation. He came back to Münster in June 2001, where his main field of research is currently functional magnetic resonance imaging and diffusion tensor imaging.

References

- Aaker DA. 1992. *Building Strong Brands*. The Free Press: New York.
- Allison RI, Uhl KP. 1964. Influence of beer brand identification on taste perception. *Journal of Marketing Research* 1(Aug): 36-39.
- Chaudhuri A, Holbrook MB. 2001. The chain of effects from brand trust and brand affect to brand performance: the role of brand loyalty. *Journal of Marketing* 65(2): 81-93.
- de Chernatony L, McDonald MH. 1998. *Creating Powerful Brands*. Butterworth Heinemann: Oxford.
- Deppe M, Schwindt W, Kugel H, Plassmann H, Kenning P. 2005a. Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *Journal of Neuroimaging* 15(2): 171-182.
- Deppe M, Schwindt W, Kugel H, Plassmann H, Kenning P, et al. (2005b). Evidence for a neural correlate of a framing effect: bias-specific activity in the ventromedial prefrontal cortex during credibility judgements. *Brain Res Bull* 67(5): 413-421.
- Erdem T, Swait J. 1998. Brand equity as a signaling phenomenon. *Journal of Consumer Psychology* 7(2): 131-159.
- Erdem T, Swait J. 2004. Brand credibility, brand consideration, and choice. *Journal of Consumer Research* 31(1): 191-198.
- Erk S, Spitzer M, Wunderlich AP, Gallo L, Walter H. 2002. Cultural objects modulate reward circuitry. *Neuroreport* 13(18): 2499-2503.
- Frackowiak RS, Friston KJ, Frith CD, Dolan RJ, Price CJ, Zeki S, Ashburner J, William Penny W. 2004. *Human Brain Function*. Academic Press: Oxford.
- Keller KL. 1993. Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing* 57(1): 1-22.
- Keller KL. 2002. *Strategic Brand Management: Building, Measuring, and Managing Brand Equity*. Prentice Hall: Upper Saddle River, NJ.
- Keller KL. 2003. Brand synthesis: the multidimensionality of brand knowledge. *Journal of Consumer Research* 29(4): 595-600.
- Kenning P, Plassmann H. 2008. How Neuroscience Can Inform Consumer Research, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* in Press.
- Koenigs M, Tranel D. 2008. Prefrontal cortex damage abolishes brand-cued changes in cola preference. *Social Cognitive and Affective Neuroscience* 3: 1-16.
- Krishnan BC, Hartline MD. 2001. Brand equity: is it more important in services? , *Journal of Services Marketing* 15(4/5): 328-3243.
- Lee N, Broderick AJ, Chamberlain L. 2007. What is "neuromarketing"? a discussion and agenda for future research. *International Journal of Psychophysiology* 63(2): 199-204.
- Lohmann H, Deppe M, Jansen A, Schwindt W, Knecht S. 2004. Task repetition can affect functional magnetic resonance imaging-based measures of language lateralization and lead to pseudoincreases in bilaterality. *Journal of Cerebral Blood Flow and Metabolism* 24(2): 179-187.
- McClure SM, Li J, Tomlin D, Cyphert KS, Montague L, Montague PR. 2004. Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44(11): 379-387.
- Paulus MP, Frank LR. (2003). Ventromedial prefrontal cortex activation is critical for preference judgments. *Neuroreport* 14(10): 1311-1315.
- Plassmann H, O'Doherty J, Rangel A. 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience* 27(37): 9984-9988.
- Shiv B, Bechara A, Levin I, Alba JW, Bettman JR, Dube L, Isen A, Mellers B, Smidts A, Grant SJ, McGraw AP. 2005. Decision neuroscience. *Marketing Letters* 16(3/4): 375-386.
- Stark CEL, Squire LR. 2001. When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences of the United States of America* 98(22): 12760-12766.
- Stigler GJ. 1961. The economics of information. *Journal of Political Economy* 69(3): 213-225.
- Stiglitz JE. 1989. Imperfect information in the product markets. In *Handbook of Industrial Organization*, Schmalensee R, Willing RD (eds). Elsevier: Amsterdam; 769-847.

In Search of Homo Economicus: Cognitive Noise and the Role of Emotion in Preference Consistency

LEONARD LEE
ON AMIR
DAN ARIELY*

Understanding the role of emotion in forming preferences is critical in helping firms choose effective marketing strategies and consumers make appropriate consumption decisions. In five experiments, participants made a set of binary product choices under conditions designed to induce different degrees of emotional decision processing. The results consistently indicate that greater reliance on emotional reactions during decision making is associated with greater preference consistency and less cognitive noise. Additionally, the results of a meta-analytical study based on data from all five experiments further show that products that elicit a stronger emotional response are more likely to yield consistent preferences.

The notion of preference consistency lies at the very foundation of understanding, predicting, and influencing consumer behavior. Most marketing activities, such as market research, new product development, marketing communications, and customer relationship management, assume consumers behave in somewhat consistent patterns. For example, if a customer indicates that he or she prefers chocolate to ginger, it is generally assumed that he or she is more likely than not to maintain such a preference in a following purchase occasion. Moreover, even people themselves like to think they are somewhat consistent decision makers (e.g., Bem 1972; Festinger 1957). On a related note, preference consistency is one of the cornerstones of *Homo*

Economicus and assumed to be characteristic of rational decision makers.

As a way to conceptualize preference consistency, consider a consumer who is faced with a series of binary choices. For each choice pair, this consumer has to evaluate the two alternatives and consider which one he or she prefers. Such a value assessment process is likely to fluctuate from case to case based on the exact information the consumer considers (e.g., the particular facts the consumer retrieves from memory), the context of the choice, as well as the particular computations that the consumer carries out; any of these process components is a potential source of “noise” and thus decision inconsistency. For example, when shopping for a new Canon digital camera, it is possible that consumers might change the aspects of the camera they focus on, the particular information they retrieve from memory, the relative importance weights they assign to the various attributes, or the process of integrating these weights.

As researchers, we often treat such inconsistencies as “noise” and use statistical inference tools that allow us to examine the data while mostly ignoring these fluctuations. Yet, such noise can convey important information about the ability of decision makers to perform good decisions, and, in particular, it can reflect their ability to conceptualize their own preferences. Moreover, from a psychometric perspective, reliability is a necessary condition for validity (Nunnally and Bernstein 1994); this property of classical test theory suggests that a decision is valid only to the extent that it is reliable. In the current work, we focus on one source of

*Leonard Lee is assistant professor of marketing, Columbia Business School, 3022 Broadway, Uris 508, New York, NY 10027 (leonardlee@columbia.edu). On Amir is assistant professor of marketing, Rady School of Management, University of California at San Diego, Pepper Canyon Hall, Room 324, 9500 Gilman Dr., MC 0093, La Jolla, CA 92093 (oamir@ucsd.edu). Dan Ariely is the James B. Duke Professor of Behavioral Economics, The Fuqua School of Business, Duke University, 1 Towerview Drive, Durham, NC 27708 (dandan@duke.edu). Correspondence: Leonard Lee. The authors would like to thank Andrew Mitchell, Michel Tuan Pham, and Duncan Simester for their comments and suggestions, Daniel Mochon, Anastasiya Pocheptosova, Danielle Suh, Xing Yuan, and SurveySpot for research assistance, and John Deighton, the associate editor, and the reviewers for their insightful comments.

John Deighton served as editor and Baba Shiv served as associate editor for this article.

Electronically published January 22, 2009

such inconsistencies or noise in decision making and its relation to preference stability: we explore whether decisions based more on emotional as opposed to cognitive processes are more prone to this kind of error. We propose that relying on one's emotional responses during decision making can promote greater preference consistency.

Dual-System Models

Prior research has established that both the emotional system and the cognitive system contribute to decision making but provide different types of inputs (Damasio 1994; Loewenstein and O'Donoghue 2004) and apply with different force within different decision environments (e.g., see McClure et al. 2004 for the different degrees of relative dominance of the two systems in choosing delayed versus immediate monetary rewards). In general, the emotional system has been characterized as being more holistic, affective, concreative, and passive, while the cognitive system is relatively more analytic, logical, abstract, and active.

This fundamental distinction between cognition and emotion is also evident across a wide array of research programs, including the emotion-cognition dual-process model (see Loewenstein and O'Donoghue 2004 for a recent review), type I versus type II processes (Kahneman and Frederick 2002), associative system versus rule-based system (Sloman 1996), nonverbal processes versus verbal processes (Paivio 1986), hot system versus cold system (Metcalfe and Mischel 1999), reflexive system versus reflective system (Lieberman et al. 2002), experiential system versus rational system (Epstein 1994, 2003), and the "approach/avoidance" versus "true/false" distinction (Zajonc 1998, 591), to name a few. Moreover, neuroscience evidence adds additional support to this distinction by demonstrating the relationship between automatic emotional response and activity in the limbic system (with the amygdala and the anterior cingulate being the prime players) and the relationship between more controlled processes and the front regions (orbital and prefrontal) of the brain (Damasio 1994; LeDoux 1996; Panksepp 2004). It should be noted that while such a dual-system conceptualization is undoubtedly an oversimplification and an imprecise representation of the complex human mind, this emotion-cognition distinction has substantial value in explaining a wide variety of human behavior.

Given this distinction between emotion and cognition and the growing discourse as to their roles in decision making, it is important to ask whether decisions that are based more on emotional inputs or more on cognitive inputs are better, or which approach individuals should adopt in decision making (e.g., should people actively suppress affective reactions when making decisions? See Pham 2007; Vohs, Baumeister, and Loewenstein 2007). Specifically, we ask the following question in the current research: to what extent does relying on one's emotional response during decision making affect the consistency of one's preferences? Following the above discussion of dual-system models and acknowledging that decisions are generated by some combination of emotion

and cognition, our question centers around the relative magnitude of reliance on emotional versus cognitive responses.

Why Might Emotions Hurt Preference Consistency?

Broadly speaking, rational behaviors (calculated, forward looking, self-controlled, value maximizing) are often attributed to the cognitive system, while irrational behaviors (myopic, transitory, lack of self-control, hyperbolic discounting, hot-cold empathy gap) are attributed to the emotional system or to a misalignment between both the emotional and the cognitive systems (Ainslie and Haslam 1992; Loewenstein and Schkade 1999; McClure et al. 2004; Mischel, Cantor and Feldman 1996; Prelec and Loewenstein 1998; Wertenbroch 2003). Given this distinction, a view appertaining to higher matching between the cognitive system and rationality and between the emotional system and irrationality is emerging (see Camerer, Loewenstein, and Prelec 2005 for a general discussion on the relationship between the automatic-controlled and affective-cognitive dichotomies).

These general associations between the cognitive system and rationality, as well as between the emotional system and irrationality, are also congruent with lay beliefs. For example, in one study, we asked a group of 16 economics or psychology PhD students to rate the extent to which they viewed a set of behaviors (long-term planning, impulsivity, self-control, etc.) as rational versus irrational, and another group of 16 economics or psychology PhD students to rate the extent to which they associated the same set of behaviors with the cognitive system or the emotional system. The results showed a high correlation ($r = 0.95$) between the irrationality-rationality ratings and the emotional-cognitive ratings. Behaviors such as long-term planning were associated with rationality and deliberate cognitive thought, while behaviors such as impulsivity were associated with irrationality and emotionality, suggesting that there is a general assumed fit between the cognitive system and rationality and between the emotional system and irrationality. Besides face validity, this intuition also has ecological validity: in a recent study of sequential financial decisions, for example, Shiv and his colleagues (2005) found that people with brain lesions focused in regions related to emotions were less impacted by past gains or losses and consequently made better, "more rational" investment decisions than normal healthy people. Based on all these results, should we then expect preference decisions that involve affective considerations to be less consistent?

Why Would Emotions Generate Greater Preference Consistency?

Despite the general associations between *Homo Economicus* and the cognitive system and between *Homo Psychologus* and the emotional system, it is possible that some aspects of behavior generally conceived as "rational" might be better generated by the emotional system, as suggested

by convergent evidence from evolutionary, social, and consumer psychology.

Evolutionary psychologists posit that the emotional system has evolved to carry out fast and accurate evaluations of important judgments and decisions. They have described emotions as a set of “programs” that have been specifically designed to solve evolutionarily recurrent situations or conditions, whether it is to fall in love, to escape from a predator, or to confront an unfaithful spouse. According to this view, such programs have the effect of activating, mobilizing, and coordinating a pool of resources, mental processes, goals, perceptual mechanisms, memory, attention, emotional expressions, and physiology toward the resolution of the adaptive problem at hand (Cosmides and Tooby 2000). Given these roles of the emotional system, it is possible that the emotional system is better attuned to consistently and reliably provide individuals with a reading of their preferences, thus creating higher speed, accuracy, and consistency.

This theorizing of the functions of emotions in evolutionary psychology corroborate Epstein’s (2003) hypothesis on the relative stability of the emotional system compared with the cognitive system: whereas the former system only changes with “repetitive or intense experience,” the latter system changes more rapidly at the “speed of thought” (Epstein 2003, 160). According to Epstein, assessments based on the emotional system tend to be more holistic, while those based on the cognitive system tend to be more analytical; as such, whereas holistic emotional processing might focus more on the gist of the target under evaluation and is hence relatively more consistent from time to time, analytical cognitive processing might be sensitive to fluctuations in any of the elements in the aforementioned preference formation process (e.g., information retrieved, decision weights, integration of information and weights) and hence susceptible to decision inconsistencies.

One source of empirical evidence for the idea that the emotional system might be associated with a higher level of consistency comes from the finding that feeling-based judgments, compared with reason-based assessments, exhibit greater interpersonal consistency (Pham et al. 2001; see also Pham 2004). In a series of experiments, Pham and his colleagues found that participants exhibited greater consensus in their feelings toward news magazine pictures and television commercials than their reason-based judgments (e.g., “This picture is good/bad” and “This picture is valuable/worthless”) of the same stimuli.

Yet further suggestive evidence for the close association between preference consistency and the emotional system comes from findings in social psychology that have challenged the value of conscious thought toward rational decision making. For instance, Wilson and his colleagues (Wilson and Schooler 1991; Wilson et al. 1993) found that deliberating the positives and negatives when evaluating hedonic experiences, such as food or art, can cause people to erroneously focus their attention on nonoptimal criteria and adversely affect the quality of their judgment. Their results showed that participants who selected jams (or post-

ers) under thoughtful deliberation later consumed less of their chosen jam (or were less likely to keep their chosen posters on their walls), compared with those whose selection was based on their immediate gut feelings. More recently, Dijksterhuis (2004) extended this work to examine the relative merits of conscious and unconscious thought in decision making. His experimental results suggest that unconscious thinkers tend to make more accurate judgments than conscious thinkers for very complex decisions. For example, in one of his studies, participants who were asked to think carefully for 4 minutes before choosing their favorite car from a given set of cars made objectively poorer choices (i.e., cars that had significantly fewer positive features) than participants who were distracted for the same amount of time with an anagram-solving task (Dijksterhuis et al. 2006, study 1; see also Nordgren and Dijksterhuis 2009).

Based on these diverse sources of evidence implicating the potentially close association between the emotional system and preference consistency, we hypothesize that a higher degree of reliance on emotional responses during decision making will generate a higher level of preference consistency.

Transitivity as a Measure of Preference Consistency

In designing an appropriate experimental paradigm to test our hypothesis, we noted the correspondence between preference consistency, the center of the current investigation, and *transitivity*—one of the two basic axioms of rational preference relations in economics which provide one set of traditional normative benchmarks for decision quality (Mas-Colell, Whinston, and Green 1995, 6). Transitivity implies that a consumer should have a well-defined preference structure, such that for any set of bundles a , b , and c , if $a \geq b$ and $b \geq c$, it must also be the case that $a \geq c$ (where \geq denotes relative preference).

In our experiments, we used transitivity (or the degree of transitivity violation) as a way to operationalize the preference consistency construct and measure the degree of (in)consistency in people’s preferences over time. Specifically, in each experiment, we took a set of P products and presented participants with all pairwise combinations of these P products (i.e., for eight products, this would mean $7 \times 8/2 = 28$ pairs) in a random order. Based on each participant’s choice pattern, we compute how many times he or she violated transitivity (e.g., $p_x \geq p_y$, $p_y \geq p_z$, and $p_z \geq p_x$, where p_x , p_y , and p_z are products within the set) as a measure of the degree of inconsistency in the participant’s preferences. Obviously, given a large number of pairwise decisions, it is inevitable that decision makers will make random errors during choice and consequently violate transitivity from time to time. Thus, simply demonstrating that such violations exist would be trivial and of negligible theoretical consequence. Instead, our interest here is to use the *degree* of intransititvity to compare whether choices are more consistent when individuals’ relative reliance on their emo-

tional responses is greater. To this end, we used different experimental manipulations aimed at invoking different degrees of relative reliance on emotional reactions when making decisions and examined whether these manipulations affect the transitivity of people's preferences, and hence their preference consistency.

At this juncture, we should note that since May's (1954) and Tversky's (1969) early demonstrations of people's systematic preference intransitivity, there has been substantial research examining the factors that can drive systematic intransitive choices. Thus far, it has been found that people's preferences can be intransitive when driven by their social context (Fishburn 1970), by the use of satisficing heuristics (Gigerenzer 2000) or noncompensatory decision rules (Tversky 1969), and by anticipated regret and counterfactuals (Loomes, Starmer, and Sugden 1991). Intransitive preferences can also occur when information for decision making is missing (Kivetz and Simonson 2000) or costly to obtain (Haines and Ratchford 1987). Given the challenge that intransitivity poses for traditional choice models, several researchers have proposed theoretical revisions to these models to accommodate preference intransitivity (e.g., Clark 1994; Fishburn 1991; Iverson and Falmagne 1985; Kim and Richter 1986; Loomes and Sugden 1983; Sopher and Gigliotti 1993).

The current work differs from these prior work in that we focus not on systematic violations of transitivity (ones that are due to biases in decision making), but on intransitivity attributed to the instability in the way decision makers consider the choices they are facing as they encounter them again and again—an intransitivity that is akin to stochasticity and that leads to diminished predictability of consumer behavior. In choice models, for example, such intransitivity would manifest as larger error terms and lower reliability and predictability.

GENERAL EXPERIMENTAL SETUP

We used the same experimental paradigm—a pairwise choice procedure—across all the five experiments we conducted. In each experiment, we first presented participants with the name, picture, and a short description of all products used in the experiment and asked them to study the products for as long as they wished (the products were electronic gadgets, such as a voice-recording key chain and a pen with a built-in FM tuner; see fig. 1 for a list of the stimuli used in the experiments). After participants had familiarized themselves with all stimuli, they were told that they would see pairs of these products and had to make a choice within each pair according to their preferences. The pairs of products were constructed by taking P products (eight products in experiments 1b and 4, and 10 products in experiments 1a, 2, and 3) and presenting participants with all pairwise combinations of these products (i.e., $P \times (P - 1)/2$; 28 pairs in experiments 1b and 4, and 45 pairs in experiments 1a, 2, and 3) in a random order.

As a measure of preference (in)consistency, we computed the number of transitivity violations participants committed

during the task for a subset of N product options, where N ranged from 3 to P . For simplicity in reporting the results, we focus on violations in the form of three-way preference cycles (e.g., $p_x \geq p_y$, $p_y \geq p_z$, and $p_z \geq p_x$; Kendall and Babington Smith 1940). If there is no error in decision making (or if it is very low), individuals will evaluate the different options in a consistent way every time and will show no (or very few) violations of transitivity in this pairwise choice paradigm. However, if individuals are prone to error in decision making, they will assess their utility very differently every time and will commit many violations of transitivity in this paradigm. Thus, we use the measure of intransitivity to capture the amount of inconsistency in evaluations, arguing that higher levels of observed intransitivity mean that the underlying process is more prone to fluctuating evaluations, inconsistencies, and random errors.

Besides the three-way-cycles measure used in the current research, other measures have been proposed for the degree of intransitivity in a sequence of pairwise choices (see Monsuur and Storcken 1975 and David 1988 for comprehensive reviews). For example, Slater's (1961) *i* counts instead the minimum number of preference reversals (on the binary level) needed to resolve all intransitivities. The experimental results in the current work are robust to the type of measure, and we adopt the three-way-cycles approach for its simplicity.

To test our research hypothesis, we used four different approaches to manipulate how much participants depend on the emotional system during choice drawing upon findings from the extant literature on the general two-system model, and compare the individual transitivity scores across the different experimental conditions. In experiments 1a and 1b, we manipulated how the stimuli were presented during the choice task (pictorial vs. verbal) based on prior research that has demonstrated that pictures trigger more emotional processing than words or symbols (Hsee and Rottenstreich 2004; see also Epstein 2003, and Lieberman et al. 2002). In experiment 2, we adapted a manipulation used by Shiv and Fedorikhin (1999) whereby color photos of the products generate greater emotional reactions than black-and-white versions of the same photos. In experiments 3 and 4, we kept the form of the choice stimuli constant, and instead manipulated participants' trust in their feelings and their cognitive capacity during choice, respectively. Prior research has demonstrated that individuals who trust their feelings more have a greater propensity to rely on their emotional system during decision making (Avnet and Pham 2007; Stephen and Pham 2008); however, choice under high cognitive load limits cognitive ability, thus generating a greater degree of reliance on emotional reactions than choice under low cognitive load (Lieberman et al. 2002; Siemer and Reisenzein 1998). Finally, we pooled together all the data from these five experiments in a meta-analytical study and examined how the nature of the products—the degree to which they generate greater affective response in consumers—can influence their contribution toward greater preference consistency. We next describe the procedures and report the re-

FIGURE 1
EXPERIMENTAL STIMULI



The FM Pen



Quick-Release Micro Light Keychain



Voice Recording Keychain with LED



Super-Bright LED Clip Light



Pagemark
Dictionary



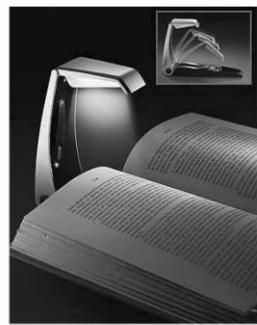
"Talking Pictures"
Photo Album for 24
pictures and 24
messages



Voice Recording Pen



LED Multi-tool



"Bright As Day!"
Daylight Spectrum Booklight



Chrome Key Organizer With Ultra-Bright Dual
LED Torch

NOTE.—The eight products in the top two rows were used in all the experiments. The additional two products in the bottom row were also used in experiments 1a, 2, and 3. (Color version available as an online enhancement.)

sults of all five experiments and the meta-analytic study in detail.

EXPERIMENT 1A: PICTURES VERSUS NAMES

Overview and Method

In the first experiment, we attempted to influence the extent to which participants relied on their emotional response during choice using a characteristic distinction of the emotional system from the cognitive system—while the emotional system is more experiential and concretive (i.e., encoding reality in concrete images, metaphors, and narratives), the cognitive system is more logical and abstract (i.e., encoding reality in abstract symbols, words, and numbers; Epstein 2003; Lieberman et al. 2002). Drawing upon this distinction, we manipulated the manner in which the visual stimuli were represented. After studying the information (name, picture, and description) of all products as described in the general procedure for an unlimited amount of time, participants were randomly assigned to one of two conditions: in the names condition, the pairs of products were presented in terms of their names, while in the pictures condition, the pairs of products were presented in terms of their pictures (Hsee and Rottenstreich 2004).

We conducted the experiment with the assistance of a commercial Web-based survey research company. A total of 534 online respondents participated and in return were entered into a sweepstakes to win a variety of prizes.

Results

A comparison of the number of three-way intransitivity errors participants made in the names and pictures conditions was performed using the unpaired sample *t*-test. Participants in the pictures condition made significantly fewer transitivity violations ($M_{pic} = 2.7$, $SD = 4.7$) than those in the names condition ($M_{names} = 4.6$, $SD = 6.3$), ($t(532) = 4.08$, $p < .001$). In addition, they spent significantly less time to make their choices ($M_{pic} = 142.0$ sec. vs. $M_{names} = 199.0$ sec.), ($t(532) = 4.32$, $p < .001$). There was, however, no difference in the amount of time participants between conditions took to study the 10 given products ($M_{pic} = 149.2$ sec. vs. $M_{names} = 143.1$ sec.), ($t(532) = .58$, $p = .56$).

Discussion

In experiment 1a, we found preliminary support for the thesis that preference consistency is associated with greater reliance on emotional responses. When the stimuli were richer in affect, participants made significantly fewer intransitivity errors. While this result is consistent with our account, it suffers from two major shortcomings: first, although pictures (vs. words) may indeed have generated a higher degree of emotional processing, they could also have facilitated participants' memory of their prior choices, hence increasing transitivity; second, because the choices that par-

ticipants made were inconsequential, it is hard to regard any of their choices as truly erroneous. We thus designed experiment 1b to address these two issues while replicating our general finding.

EXPERIMENT 1B: CONSEQUENTIAL PICTURES AND/OR NAMES

Overview and Method

Experiment 1b replicated experiment 1a in a laboratory setting with two important differences. First, in addition to the names condition and the pictures condition, we added a third condition in which both the names and pictures of the product options were available at the choice stage. Arguably, this new (combined) condition presents more information than pictures or words alone; thus, if the ease-of-recall account were sufficient to explain our results in experiment 1a, participants ought to be most transitive in this combined condition. However, if our hypothesis is true, then the inclusion of the names of the products (with their pictures) should activate semantic cognitive processing and render this condition no different from the names-only condition, given that we hypothesize that it is the cognitive processing that adds noise to the decision process. Second, choices were consequential as participants had the chance to receive one of their product choices, hence making the experiment incentive compatible. Specifically, participants were told that at the end of the experiment, one of the 28 pairs of products would be picked at random and they would be entered into a lottery to win the product they picked in this pair.

A total of 75 students recruited at the MIT Stratton student center participated in this experiment and were randomly assigned to one of the three conditions: names-only, pictures-only, and names-and-pictures. They were each paid \$1 for their participation and were told they would be entered into a lottery to win one of the products they chose.

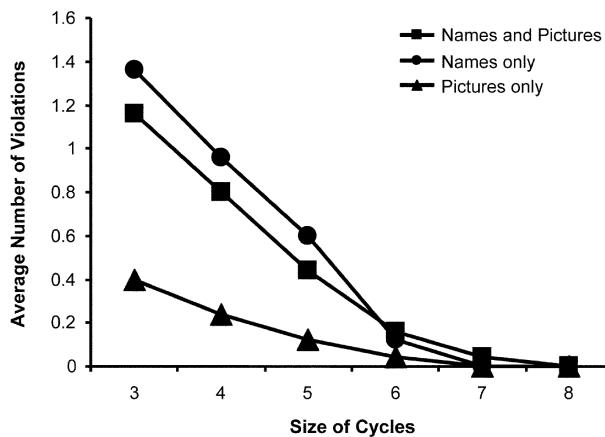
Results

The results paralleled closely those of experiment 1a: participants in the pictures-only condition who chose between products presented in the form of pictures made significantly fewer intransitivity errors ($M_{pic-only} = .4$, $SD = 1.0$) than those in the names-only condition who chose between products presented in the form of names ($M_{name-only} = 1.4$, $SD = 1.7$), ($t(48) = 2.51$, $p = .02$). In addition, when the product pairs were displayed in terms of both types of information (picture and name), the number of violations ($M_{name-&pic} = 1.2$, $SD = 1.4$) were similar to those in the names-only condition ($t(48) = .46$, $p = .65$), but greater than those in the pictures-only condition ($t(48) = 2.20$, $p = .03$). (This result pattern continued to hold when we tested for differences in the number of larger cycles across conditions, as shown in fig. 2.)

Overall, these results (overall ANOVA: $F(2, 72) = 3.37$, $p = .04$) suggest that alternative accounts that speak to the

FIGURE 2

COMPARISON OF THE AVERAGE NUMBER OF INTRANSITIVITY CYCLES (EXPERIMENT 1B)



product pictures' being a better memory retrieval cue may not be sufficient to explain our results. The availability of product names (in addition to the pictures) was sufficient to deteriorate choice consistency. Given that different modes of stimuli presentation generate different degrees of emotional/cognitive processing, the current results suggest that preference consistency is more closely associated with affective processing than more controlled cognitive processing. Furthermore, there was no significant difference across the three conditions in the amount of time participants took to study the products prior to choice ($M_{\text{pic-only}} = 113.7$ sec. vs. $M_{\text{name-only}} = 114.7$ sec. vs. $M_{\text{name-&-pic}} = 104.2$ sec.), ($F(2, 72) = .14$, $p = .87$) or the amount of time they took to choose their preferred products ($M_{\text{pic-only}} = 100.5$ sec. vs. $M_{\text{name-only}} = 118.1$ sec. vs. $M_{\text{pic-only}} = 98.2$ sec.), ($F(2, 72) = 1.5$, $p = .23$).

Discussion

Consistent with the hypothesis that greater reliance on emotional responses contributes toward a higher degree of preference consistency, the results of this experiment show that manipulations (i.e., color images of products) designed to tap automatic emotional processes to a greater extent than controlled cognitive processes can generate higher levels of preference consistency than manipulations (i.e., names of products) designed to tap cognitive processes to a greater extent than emotional processes. While the levels of preference consistency differed across the experimental conditions, it is also worthwhile to obtain a sense of the overall magnitude of consistency observed in this experiment. One approach is to consider the following: if the participants had chosen at random in this experiment, they would have made an average of about 14 intransitivity cycles (computed based on a simulation of 10,000 participants who made random choices). Clearly, this quantity is much higher than anything

we found; in comparison with this benchmark, the overall level of intransitivity in experiment 1b was rather low.

The manipulation of pictures versus names used in the first two experiments is consistent with our proposed account, but it is also clear that there are many differences between pictorial information and textual information. Though we have managed to rule out some potential alternative explanations (particularly with the combined names-and-pictures condition), a different approach for testing the main hypothesis would provide useful converging evidence. Experiment 2 was designed to test the hypothesis using a different manipulation of the extent to which different stimuli-display modes evoke emotional responses.

EXPERIMENT 2: COLOR VERSUS BLACK-AND-WHITE PICTURES

Overview and Method

Previous research has shown that the vividness of pictorial information influences the degree of emotionality experienced by consumers (Loewenstein 1996; Mischel and Moore 1973; Shiv and Fedorikhin 1999). In experiment 2, we used a similar approach to manipulate the vividness and emotionality of the product options by presenting participants either color pictures or black-and-white (B&W) pictures of the products in the choice task.

A total of 88 students recruited at Princeton participated in this lab experiment. After participants had familiarized themselves with the 10 products (in the same unhindered manner as before), each picture being represented by a name, a picture, and a short description, they proceeded to make a sequence of 45 pairwise choices. Participants were randomly assigned to one of two conditions: half the participants were presented with the names and color pictures of the 10 products; the other half were presented with the names and pictures of the same products, but the pictures were in B&W. To examine whether the different presentation modes of the pictures affect the participants' perception of how much information they were obtaining from the pictures, at the end of the choice task, we asked all participants to complete a postchoice survey in which they had to rate, on a scale of 1 (not at all) to 7 (very well), how well they thought each of the given pictures adequately represented the corresponding product described. As in experiment 1b, participants were told that they would be entered into a lottery to win one of the products they chose. Based on the previous experiments, we hypothesize that participants who were presented with the color pictures of the product options would exhibit greater preference consistency and more transitive choices than those in the B&W condition.

Results

An analysis of the number of intransitivity errors participants made between the two conditions revealed that participants who saw the black-and-white pictures during the choice task committed twice as many transitivity violations

($M_{b\&w} = 2.2$, $SD = 2.74$) as participants who were presented with the color pictures instead ($M_{color} = 1.1$, $SD = 1.86$), ($t(86) = 2.05$, $p = .04$). The participants in the B&W condition also took marginally more time both to study the products ($M_{b\&w} = 123$ sec. vs. $M_{color} = 98.5$ sec.), ($t(86) = 1.92$, $p = .06$) and to make their choices ($M_{b\&w} = 120$ sec. vs. $M_{color} = 109.7$ sec.), ($t(86) = 1.66$, $p = .10$) than those in the color condition. Furthermore, the results of the postchoice representation survey revealed that participants across the two conditions did not report any significant difference in how adequately they thought the pictures represented the products (for each of the 10 products: $p = .25$ to $.95$; for the total ratings of all 10 products: ($M_{b\&w} = 49.6$, $SD = 8.7$ vs. $M_{color} = 49.1$ sec., $SD = 8.9$), ($t(86) = .29$, $p = .77$)).

Discussion

The results of this experiment continue to implicate the close association between relying on emotional reactions and preference consistency. In this experiment, instead of manipulating the type of information participants saw during choice (names vs. pictures), we focused on eliciting different degrees of emotional reactions during decision making by manipulating the vividness of the stimuli presented to participants. The results of the postchoice survey also suggest that this difference in preference consistency cannot be adequately explained by any potential difference in the perceived amount of product information obtained through the different presentation modes. Participants did not seem to discern any significant difference in objective product information between the two conditions; rather, the color versus B&W pictures elicited different degrees of emotional reactions in participants. This finding also renders alternative accounts such as differential memory recall or differential processing fluency less likely.

EXPERIMENT 3: TRUST IN FEELINGS

Overview and Method

So far, we have manipulated the manner in which the choice stimuli were presented to participants in order to evoke different degrees of emotional responses in participants. One way to test our hypothesis more directly is to explicitly manipulate the cognitive state of participants when making decisions while keeping the choice stimuli constant across conditions. To this end, we used a method by Avnet and Pham (2007; see also Stephen and Pham 2008) to manipulate participants' situational trust in their emotions and thus their reliance on affect as information during decision making. If reliance on emotional reactions indeed generates greater preference consistency, then participants who trust their feelings more should exhibit more consistent preferences (regardless of the presentation format of choice stimuli) and make fewer intransitivity errors than those who trust their feelings less during decision making. Conversely, participants who trust their feelings less might display different

degrees of preference consistency depending on how much emotional response the choice stimuli generate.

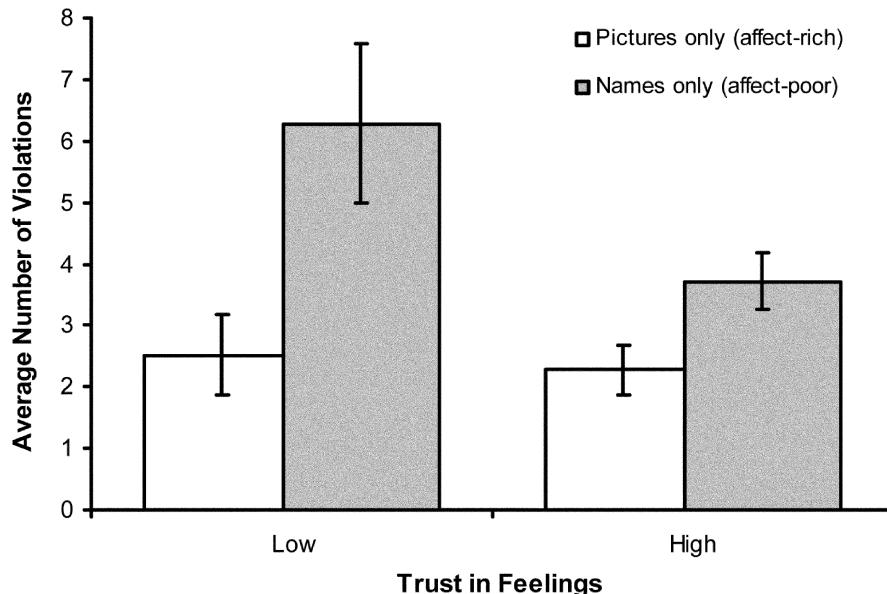
This experiment was conducted as two purportedly separate online studies. (Participants' response to a postexperiment survey question indicated that they were not aware of the relationship between these two online studies.) In the first study, participants were asked to describe either two (high-trust) or 10 (low-trust) past situations in which they trusted their feelings to make a decision and it turned out to be the right decision. This manipulation is based on the premise that participants who had to describe two situations would find the task relatively easy, whereas those who had to describe 10 situations would find the task relatively difficult (Schwarz et al. 1991); this experienced ease or difficulty of retrieving two versus 10 instances, respectively, would then result in respondents' perceiving that such instances are common versus uncommon, thus generating higher (two) versus lower (10) trust in their feelings when making subsequent decisions. In the second study, participants were given the same general two-stage binary choice task as used in all earlier experiments. However, as in experiment 1a, half the participants saw the products presented in terms of pictures during the choice stage, while the other half saw the products presented in terms of names. We predict that while high-trust participants would demonstrate a similar degree of choice consistency regardless of whether the products were presented in terms of names or pictures, low-trust participants would be more affected by the presentation format of the products: in particular, when the products were represented in terms of pictures instead of names, low-trust participants would be more likely to rely on their feelings when choosing between the products and demonstrate greater consistency.

A total of 208 individuals recruited from an online panel participated in this experiment. (Six participants were excluded from the analysis because the amount of time they took to complete the task was more than three SDs away from the average time based on a logarithmic transformation of the completion time; including these participants within the data set, however, did not change the general result pattern.) Each participant was randomly assigned to one of the four conditions in the 2 (trust in feelings: high vs. low) \times 2 (product representation: pictures vs. names) between-subjects design and stood a chance to win one of the chosen products in a lottery after the experiment.

Results

The number of intransitivity errors that participants made in the binary choice task was submitted to a two-factorial (trust in feelings \times product representation) ANOVA. As shown in figure 3, the results revealed a statistically significant main effect of product representation ($F(1, 198) = 13.92$; $p < .001$), a significant main effect of trust in feelings ($F(1, 198) = 4.04$, $p = .05$), and a marginally significant interaction between the two independent factors ($F(1, 198) = 2.73$, $p < .10$). Most central to the objective of this experiment, planned contrasts revealed that whereas high-trust par-

FIGURE 3
COMPARISON OF THE AVERAGE NUMBER OF INTRANSITIVITY CYCLES (EXPERIMENT 3)



NOTE.—Error bars denote standard errors.

ticipants did not differ significantly in the number of intransitivity errors that they made regardless of whether the products were presented in terms of pictures ($M = 2.27$, $SD = 3.06$) or names ($M = 3.72$, $SD = 3.55$), ($F(1, 198) = 2.54$, $p = .11$), low-trust participants made significantly fewer intransitivity errors when the products were presented in terms of pictures ($M = 2.52$, $SD = 4.36$) compared with names ($M = 6.28$, $SD = 8.19$), ($F(1, 198) = 13.92$, $p < .001$).

In contrast, ANOVAs comparing the amount of time participants across conditions took to make their choices revealed only a significant main effect of product representation ($F(1, 198) = 11.83$, $p < .001$), with participants taking less time to complete the task when pictures ($M = 111.88$ sec., $SD = 45.21$ sec.) were presented instead of names ($M = 138.44$ sec., $SD = 57.80$ sec.). Neither the main effect of trust in feelings or the interaction effect between the two independent factors on decision time was significant (both p 's $> .44$). The amount of time participants took to study the products before choice was also not significant across conditions (all p 's $> .37$).

Discussion

In addition to conceptually replicating our previous findings involving different ways in which the choice stimuli were represented, experiment 3 adds an important insight into the processes underlying preference consistency: by encouraging trust in feelings, we can induce greater emotional processing and increase preference consistency, es-

pecially when products are not presented in a way that naturally engenders affective processing. This approach seems to contradict potential lay advice to inhibit one's feelings in order to generate more consistent preferences.

One alternative explanation to the results is that, compared with participants who had to recall only two previous incidents, those who had to recall 10 previous incidents experienced greater cognitive depletion and hence were less careful with their decisions, thereby making more intransitivity errors. This account seems inadequate at explaining the observed transitivity pattern since participants who were asked to recall 10 previous incidents did not take a significantly different amount of time on average to make their choices compared with those asked to recall only two previous incidents ($p = .55$). Nonetheless, in the next experiment, we directly manipulated participants' cognitive capacity and hence their *relative* reliance on their emotional processing.

EXPERIMENT 4: COGNITIVE CAPACITY

Overview and Method

In this experiment, we used the same general procedure as in the earlier experiments, first presenting participants with the names, pictures, and descriptions of the products for study and then asking them to choose within pairs of these products (represented by both their names and pictures). As in experiment 3, the choice stimuli did not vary in how they were presented to participants across conditions.

Instead, to induce different degrees of relative reliance on emotional reactions, we used a cognitive load manipulation that has been widely adopted in the psychology literature (Shiv and Fedorikhin 1999; Trope and Alfieri 1997; see also Gilbert, Pelham, and Krull 1988): half the participants in the experiment were asked to memorize a three-digit code (low-load condition) during the choice task, while the other half were asked to memorize a 10-digit code (high-load condition). The pretext for the code recall was that participants had to reproduce the correct code at the end of the choice task to enter a lottery to win one of the products they had selected.

Based on prior research (Lieberman et al. 2002; Siemer and Reisenzein 1998), we expected participants in the high-load condition—whose cognitive capacities were constrained by the requirement to memorize a long numeric code—to rely *more* on their emotional responses when choosing their preferred products. Thus, if preference consistency is indeed associated with a greater reliance on emotional reactions, then participants in the high-load condition should make fewer intransitivity errors than those in the low-load condition. However, if preference consistency is associated with cognitive processing instead, then we would see the opposite pattern of results.

Forty students at MIT participated in this study in exchange for the opportunity to win one of the products they chose in the choice task. The students were randomly assigned to one of the two experimental conditions.

Results

An unpaired-sample *t*-test revealed that the high-load participants committed significantly fewer transitivity violations ($M_{\text{high}} = .7$, $SD = 1.03$) than the low-load participants ($M_{\text{low}} = 1.9$, $SD = 2.49$), ($t(38) = 1.99$, $p = .05$). There was, however, no significant difference between the two conditions in the amount of time participants took to study the products prior to choice ($M_{\text{low}} = 89$ sec. vs. $M_{\text{high}} = 78$ sec.), ($t(38) = .46$, $p = .65$), nor the amount of time they took to choose the products ($M_{\text{low}} = 85.3$ sec. vs. $M_{\text{high}} = 85.3$ sec.), ($t(38) < .01$, $p > .99$).

Discussion

Consistent with our hypothesis, we found that participants whose cognitive capacity was constrained were *more* consistent in their choices. In accordance with prior research, participants in the high-load condition had to rely more on emotional as opposed to cognitive processes than their low-load counterparts when selecting between each pair of products. That participants in the high-load condition made fewer intransitivity errors provides further support for our claim that preference transitivity is associated more closely with reliance on emotional reactions than with deliberate cognitive thought. Put differently, cognitive processes are more prone to stochastic noise. Note that an alternative account based on cognitive depletion, as discussed in experiment 3, can be effectively ruled out here given that participants in

the high-load condition, despite being cognitively more depleted than participants in the low-load condition, actually made fewer, not more (as the alternative account would predict), intransitivity errors.

A META-ANALYTICAL STUDY: THE ROLE OF PRODUCT TYPES

Our data from the five experiments allow for an even stronger test of our hypothesis. If indeed, as we claim, the choice inconsistencies we observe in our experiments were driven by the cognitive system's greater susceptibility to decisional noise (or "cognitive noise") compared with the emotional system, then this effect should be larger for products that generate little emotional reaction than for products that generate greater emotional response. In other words, if we could classify the products themselves based on the extent to which they generate emotional responses in people, then we expect those products that elicit a greater emotional reaction to generate greater preference consistency.

To test this hypothesis in the most conservative manner, we took all product choices (across all of our experiments) and asked the following question: are products that generate a stronger emotional reaction less likely to belong to intransitivity cycles than products that generate a more cognitive response? To answer this question, we presented the same product information (i.e., names and pictures) that was available to all of our experimental participants to an independent group of 30 university students and asked them to rate on a 10-point scale the extent to which each of the products was "functional/useful" and the extent to which the product was "exciting/cool." The correlation between the two scores for each product was $r = .89$ ($p = .0004$). Nevertheless, for each product, the average difference between these two measures represented the extent to which the product elicited a greater emotional versus cognitive reaction. We then analyzed whether being "more exciting than useful" (METU) indeed predicted a smaller likelihood for a product to belong to an intransitivity cycle.

A random-effects probit regression analysis of the likelihood of a product to belong to an intransitivity cycle on the product's METU measure as an independent variable, and controlling for individual heterogeneity, revealed that the stronger is a product's emotional relative to cognitive appeal (i.e., greater METU), the *less* likely is the product to belong to an intransitivity cycle ($p < .001$; see table 1, model I). (Adding interaction terms between product attributes and specific within-experiment manipulations into the regression equation did not produce any reliable effects; the same basic result pattern also held when alternative methods were used to account for participant heterogeneity.) Interestingly, although decision times across conditions (within the individual experiments) were not always significantly different and could not fully explain the observed difference in preference consistency between conditions, consistent with a theory that making emotionally-based choices is somewhat faster, adding decision time as a predictor in the

TABLE 1**RESULTS OF INDIVIDUAL-SPECIFIC RANDOM-EFFECTS PROBIT REGRESSIONS PREDICTING INTRANSITIVITY**

Predictor	Model I	Model II
Intercept	1.508** (.038)	1.510** (.039)
METU	-.078** (.015)	-.077** (.015)
Decision time		7.45e-07* (3.63e-07)
n / # of groups	40,300 / 939	40,300 / 939

NOTE.—Standard errors are presented in parentheses below parameter estimates.

* $p < .05$.

** $p < .001$.

regression (see table 1, model II) revealed a small yet significant effect such that longer decision times for a product were associated with a greater likelihood of transitivity violation ($p < .05$). However, differences in decision time did not mediate the role of product emotionality in fostering transitivity, as the METU independent variable remained statistically significant after controlling for decision time ($p < .001$).

We thus find that not only did more “emotion-laden” conditions within experiments resulted in greater choice consistency as shown in the previous five experiments, preferences for products that evoke a stronger emotional reaction across conditions also tend to be more consistent. This finding lends further support to the differential roles of emotional processing and cognitive processing in generating consistent preferences. Again, we find that with greater reliance on emotional response comes a higher level of transitivity and greater preference consistency.

GENERAL DISCUSSION

We arrive at the truth, not by the reason only, but also by the heart. (Blaise Pascal)

As the canonical symbol of rational decision making, *Homo Economicus* has generally been depicted as a superrational, self-interested breed that possesses immense foresight and cognitive abilities (and perhaps, consequently, an oversized and active cortical system) but at the same time, “devoid of emotions.” In this work, we investigated one important property of *Homo Economicus*—transitivity—and its relation to our emotional system and cognitive system. Our examination of transitivity in the current work was not aimed to test the validity of rational choice models or economic theory. Instead, we used this very central concept in economics as an apparatus to examine the consistency in which individuals decide among their choices. Beyond its theoretical value, consistency is at the core of marketers’ ability to forecast and predict consumer behavior.

The results of five experiments—in which we manipulated the visual form of the choice stimuli and the mental

state of the decision maker—consistently demonstrate that the predictability of behavior relies more on emotion than what common conceptions of decision making might suggest. Using a pairwise choice task and different experimental manipulations designed to activate different degrees of reliance on emotional versus cognitive reactions, we found that participants’ preferences were more consistent and less susceptible to cognitive noise when they chose between products presented in the more affective mode of pictures instead of the less affective mode of names (experiments 1a, 1b, and 3); when they chose between products presented in the more affective mode of color photos instead of the less affective mode of black-and-white photos (experiment 2); when they made their choices with a higher degree of trust in their feelings (experiment 3); and under higher cognitive load (experiment 4), as well as when they were making choices among products that naturally engender greater emotional than cognitive response (product meta-analysis). Together, these results imply that preference consistency is greatly benefited by affective responses. From a methodological perspective, these results also highlight the importance of examining the consistency of individuals’ choices over time as an indication of decision quality, rather than treating choice variations and inconsistencies merely as “noise.”

The ongoing discourse regarding the role of emotions in decision making presents a complex set of evidence pointing both for and against their merits in decision making. A closer analysis of the various sets of findings, including the results presented here, suggests that there are some situations in which relying on one’s emotions may be the right strategy, but other situations in which such reliance may be detrimental for decision making (Vohs et al. 2007). For instance, there are various degrees of emotional reactions, ranging from attention toward affective information (even automatically) to violent mood swings, which may have different effects on the quality of the decision making process. In the same vein, Baumeister, DeWall, and Zhang (2007) distinguished between “automatic affect” (i.e., quick reactions of liking and disliking) and “conscious emotions” (i.e., complete emotional experiences imbued with conscious feelings and cognitive interpretations). (See also Camerer, Loewenstein, and Prelec 2005 for a discussion of the distinction between automatic emotions and controlled emotions.) In this work, we focus on the former rather than the latter type of emotional processes; nevertheless, this distinction is important here because it cautions us when generalizing our results to the realm of very conscious or very strong emotional states.

Our experimental findings also extend the stream of research on preference transitivity: in addition to other antecedents that have already been identified, preference transitivity can be impaired by too much deliberate cognitive thinking. Our results are also consistent with the evolutionary account that our emotional processes might have been adapted to perform common and important tasks effectively and efficiently (Cosmides and Tooby 2000; Damasio 1994).

To the extent that transitive preferences are objectively “better” and more optimal than intransitive preferences, our results join other prior work that demonstrate the positive roles emotions play in decision making (Bechara et al. 1997; Damasio 1994; Davidson et al. 2000; Dijksterhuis 2004; Loewenstein and Lerner 2002; LeDoux 1996; Peters and Slovic 2000; Pham et al. 2001; Wilson and Schooler 1991). In particular, whereas Pham and his colleagues (2001) demonstrated that feeling-based processes (compared with reason-based processes) lead to more stable and consistent judgments across individuals, our results show that emotional processes can also contribute toward greater preference stability and consistency within individuals.

Alternative Accounts

Alternative accounts bear the burden of explaining why we observed greater preference consistency in conditions that induced more emotional and/or less cognitive processing across all the experiments and analyses presented. One such possible account for our experimental results in general is that different experimental manipulations, or the different circumstances under which participants had to make their product choices, changed the type of decision strategies participants used, and that some of these strategies somehow increase internal preference consistency. It should be noted that to the extent that these latter strategies involve more emotional processing (and hence less cognitive processing), they are essentially the rationale for our argument. Therefore, for this alternative account to be valid, the manipulations must have led to different types (and not degrees) of cognitive processing.

For example, in experiment 4 where we manipulated the degree of cognitive capacity under which participants had to choose their preferred products, one might argue that participants under higher cognitive load could have used a simplifying noncompensatory decision strategy, which in turn could have resulted in the greater degree of transitivity observed. However, Tversky’s (1969) transitivity findings with gamble choice would lead us to predict the opposite result: if a simplifying or lexicographic decision strategy was indeed used by participants under high cognitive load, then we would expect participants under low cognitive load to have greater attentional capacity to use a more optimizing, more compensatory strategy in making their choices, and thus be more, not less, consistent in their choices! (See also Gigerenzer 2000 for examples of intransitivity caused by the use of satisficing, noncompensatory decision strategies.) Moreover, the timing results that we found in this experiment—that there was no significant difference in how long participants took to choose across both conditions—further challenge the validity and adequacy of this alternative account in explaining our results. Furthermore, we also designed the experimental procedure such that participants were provided with full information of all products at the outset and were given as much time to familiarize themselves with the products as they wanted before being assigned to one of the experimental conditions. Thus, we do

not expect participants to make vastly different choices among the products in any of the experiments (which had identical product assortments).

To test this alternative account more directly using the choice data from the experiments, we examined whether there was a change in participants’ overall preference ordering among the given products across conditions, assuming that a change in decision strategy would lead to a change in preference structure. We can infer a participant’s preference order by computing his or her Kendall score for each product (i.e., the number of times the participant chooses the product in preference to the other products in the set), and rank order his or her Kendall scores for all the products (Cook and Kress 1992). (In other words, there is a set of P Kendall scores for each participant, where P is the number of products in the given set.) To compare participants’ preference structure between conditions in each experiment, we submitted participants’ Kendall scores to a MANOVA, using the specific experimental manipulation as the independent factor. The MANOVA results (using Wilks’s lambda) revealed no significant difference in product preference order across conditions in any of the experiments.

Together, these results support our proposed account as a valid parsimonious explanation for the experimental results—participants who rely on more intuitive, emotional processing rather than deliberative, cognitive processing tend to exhibit greater preference consistency.

Future Research

In this research, we measured preference consistency by computing the number of transitivity violations among all pairwise choices of a product set. There are no doubt other ways to measure consistency, and it would be worthwhile to design and conduct further experiments based on other measures of consistency. Other potential factors, such as choice context (e.g., Amir and Levav 2008) or the consideration of specific product attributes (e.g., Lee, Bertini, and Ariely 2008) that could influence the degree of emotional or cognitive processing and in turn preference consistency can also be further explored.

One important way to extend the results is to implement a design with greater temporal distance between consecutive pairwise choices, for example, getting participants to make one pairwise choice a day over an extended period of time, and examining whether the same transitivity patterns we observed across different emotional versus cognitive decision making scenarios persist. Another possible area for future research stems from an application of Piaget’s (1969) theory of cognitive development in developmental psychology. Piaget proposed four stages of cognitive development, which includes a stage (III) for concrete operations (approximately at ages 7–11) when children master logic and develop “rational” thinking. Piaget’s work suggests that age could be a factor that moderates the degree of transitivity of an individual’s choices. As such, it would be worthwhile and interesting to examine how adults compare with children in terms of their degrees of preference consistency. At a

more rudimentary (neural) level, while our experimental results point toward the close association between the emotional system and preference consistency, further experiments involving the direct examination of individuals' neural activities during choice and decision making under varying conditions are imperative if we are to identify the precise neural substrates associated with consistent preferences.

Practical Implications

Our results suggest that marketers who study consumer preferences may improve their methods by using affect-rich stimuli. For example, a conjoint analysis that includes pictures in addition to descriptions in its choice stimuli may do a far better job at tapping the more consistent components of consumer preferences. Moreover, the emotional system's contribution to stable preferences suggested by our findings points toward yet another advantage in emotional approaches to persuasion: it may be important to tap consumers' emotional systems when assessing their consumption satisfaction as this approach might yield better predictions of their future choices. Finally, our results have one other practical implication if we were to consider an important difference between brick and mortar shopping and Internet shopping: the former affords consumers a richer affective shopping experience and may lead to greater preference consistency over time. As our results suggest, the degree to which a shopping Web site provides rich affective cues might have a great influence on consumer preference consistency.

For the consumer, contrary to lay perceptions, attending to one's emotional responses may prove to be very valuable in understanding one's inherent preferences (Simonson 2008). It is possible consumers would be much happier with choices based more on their emotional reaction. For example, if one buys a house and relies on very cognitive attributes such as resale value, one may not be as happy actually living in it, as opposed to a person who attends to his or her emotional reaction to the house prior to purchasing it. Indeed, our results suggest that the heart can very well serve as a more reliable compass to greater long-term happiness than pure reason.

REFERENCES

- Ainslie, George and Nick Haslam (1992), "Hyperbolic Discounting," in *Choice over Time*, ed. George Loewenstein and Jon Elster, New York: Russell Sage, 57–92.
- Amir, On and Jonathan Levav (2008), "Choice Construction versus Preference Construction: The Instability of Preferences Learned in Context," *Journal of Marketing Research*, 45, 145–58.
- Avnet, Tamar and Pham, Michel T. (2007), "Metacognitive and Nonmetacognitive Reliance on Affect as Information in Judgment," working paper, Yeshiva University, New York.
- Baumeister, Roy F., C. Nathan DeWall, and Liqing Zhang (2007), "Do Emotions Hinder or Improve the Decision Making Process?" in Vohs et al. 2007, 11–31.
- Bechara, Antonine, Hanna Damasio, Daniel Tranel, and Antonio R. Damasio (1997), "Deciding Advantageously before Knowing the Advantageous Strategy," *Science*, 275, 1293–94.
- Bem, Daryl J. (1972), "Self-Perception Theory," in *Advances in Experimental Social Psychology*, Vol. 6, ed. Leonard Berkowitz, New York: Academic Press, 1–62.
- Camerer, Colin, George Loewenstein, and Drazen Prelec (2005), "Neuroeconomics: How Neuroscience Can Inform Economics," *Journal of Economic Literature*, 43 (1), 9–64.
- Clark, Stephen A. (1994), "Toward a Unified Model of Weakly Rational Choice," *Journal of Mathematical Psychology*, 38, 73–85.
- Cook, Wade D. and Moshe Kress (1992), *Ordinal Information and Preference Structures: Decision Models and Applications*, Upper Saddle River, NJ: Prentice Hall.
- Cosmides, Leda and John Tooby (2000), "Evolutionary Psychology and the Emotions," in *Handbook of Emotions*, ed. Michael Lewis and Jeannette M. Haviland-Jones, New York: Guilford, 91–115.
- Damasio, Antonio R. (1994), *Descartes' Error*, New York: Grosset/Putnam.
- David, H. A. (1988), *The Method of Paired Comparisons*, London: Griffin.
- Davidson, Richard J., Daren C. Jackson, and Ned H. Kalin (2000), "Emotion, Plasticity, Context, and Regulation: Perspectives from Affective Neuroscience," *Psychological Bulletin*, 126 (6), 890–909.
- Dijksterhuis, Ap (2004), "Think Different: The Merits of Unconscious Thought in Preference Development and Decision Making," *Journal of Personality and Social Psychology*, 87 (5), 586–98.
- Dijksterhuis, Ap, Maarten W. Bos, Loran F. Nordgren, and Rick B. van Baaren (2006), "On Making the Right Choice: The Deliberation-without-Attention Effect," *Science*, 311, 1005–7.
- Epstein, Seymour (1994), "Integration of the Cognitive and the Psychodynamic Unconscious," *American Psychologist*, 49, 709–24.
- (2003), "Cognitive-Experiential Self-Theory of Personality," in *Handbook of Psychology*, Vol. 5, *Personality and Social Psychology*, ed. Theodore Millon, Melvin J. Lerner, and Irving B. Weiner, Hoboken, NJ: Wiley, 159–84.
- Festinger, Leon (1957), *A Theory of Cognitive Dissonance*, Evanston, IL: Row, Peterson.
- Fishburn, Peter C. (1970), "The Irrationality of Transitivity in Social Choice," *Behavioral Science*, 15, 119–23.
- (1991), "Nontransitive Preferences in Decision Theory," *Journal of Risk and Uncertainty*, 4 (2), 113–34.
- Gigerenzer, Gerd (2000), *Adaptive Thinking: Rationality in the Real World*, New York: Oxford University Press.
- Gilbert, Daniel T., Brett W. Pelham, and Douglas S. Krull (1988), "On Cognitive Busyness: When Person Perceivers Meet Persons Perceived," *Journal of Personality and Social Psychology*, 54 (May), 733–40.
- Haines, George H., Jr. and Brian T. Ratchford (1987), "A Theory of How Intransitive Consumers Make Decisions," *Journal of Economic Psychology*, 8, 273–98.
- Hsee, Christopher K. and Yuval Rottenstreich (2004), "Music, Pandas, and Muggers: On the Affective Psychology of Value," *Journal of Experimental Psychology: General*, 133 (1), 23–30.

- Iverson, Geoffrey and J. C. Falmagne (1985), "Statistical Issues in Measurement," *Mathematical Social Sciences*, 10, 131–53.
- Kahneman, Daniel and Shane Frederick (2002), "Representativeness Revisited: Attribution Substitution in Intuitive Judgment," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. Tom Gilovich, Dale Griffin, and Daniel Kahneman, New York: Cambridge University Press, 49–81.
- Kendall, M. G. and B. Babington Smith (1940), "On the Method of Paired Comparisons," *Biometrika*, 31, 324–45.
- Kim, Taesung and Marcel L. Richter (1986), "Nontransitive-Non-total Consumer Theory," *Journal of Economic Theory*, 38 (2), 324–63.
- Kivetz, Ran and Itamar Simonson (2000), "The Effects of Incomplete Information on Consumer Choice," *Journal of Marketing Research*, 37 (4), 427–48.
- LeDoux, Joseph E. (1996), *The Emotional Brain: Mysterious Underpinnings of Emotional Life*, New York: Simon & Schuster.
- Lee, Leonard, Marco Bertini, and Dan Ariely (2008), "Money Muddles Thinking: The Effect of Price on Preference Consistency," working paper, Columbia University.
- Lieberman, Matthew D., Ruth Gaunt, Daniel T. Gilbert, and Yaacov Trope (2002), "Reflexion and Reflection: A Social Cognitive Neuroscience Approach to Attributional Inference," in *Advances in Experimental Social Psychology*, Vol. 34, ed. Mark P. Zanna, San Diego: Academic Press, 199–249.
- Loewenstein, George (1996), "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Process*, 65 (March), 272–92.
- Loewenstein, George and Jennifer Lerner (2002), "The Role of Emotion in Decision Making," in *Handbook of Affective Science*, ed. Richard J. Davidson, Klaus R. Scherer, and H. Hill Goldsmith, Oxford: Oxford University Press, 619–42.
- Loewenstein, George and Ted O'Donoghue (2004), "Animal Spirits: Affective and Deliberative Processes in Economic Behavior," working paper, Carnegie Mellon University.
- Loewenstein, George and David Schkade (1999), "Wouldn't It Be Nice? Predicting Future Feelings," in *Well-Being: The Foundations of Hedonic Psychology*, ed. Daniel Kahneman, Ed Diener, and Norbert Schwarz, New York: Russell Sage, 85–108.
- Loomes, Graham, Chris Starmer, and Robert Sugden (1991), "Observing Violations of Transitivity by Experimental Methods," *Econometrica*, 59 (2), 425–39.
- Loomes, Graham and Robert Sugden (1983), "A Rationale for Preference Reversals," *American Economic Review*, 73, 428–32.
- Mas-Colell, Andreu, Michael Dennis Whinston, and Jerry R. Green (1995), *Microeconomic Theory*, New York: Oxford University Press.
- May, Kenneth O. (1954), "Intransitity, Utility, and the Aggregation of Preference Patterns," *Econometrica*, 22, 1–13.
- McClure, Samuel M., David I. Laibson, George Loewenstein, and Jonathan D. Cohen (2004), "Separate Neural Systems Value Immediate and Delayed Monetary Rewards," *Science*, 206, 503–7.
- Metcalfe, Janet and Walter Mischel (1999), "A Hot/Cool-System Analysis of Delay of Gratification: Dynamics of Willpower," *Psychological Review*, 106 (1), 3–19.
- Mischel, Walter, Nancy Cantor, and Scott Feldman (1996), "Principles of Self-Regulation: The Nature of Willpower and Self-Control," in *Social Psychology: Handbook of Basic Principles*, ed. E. T. Higgins and A. W. Kruglanski, New York: Guilford, 329–61.
- Mischel, Walter and Bert Moore (1973), "Effects of Attention to Symbolically Presented Rewards on Self-Control," *Journal of Personality and Social Psychology*, 28, 172–79.
- Monsuur, Herman and Ton Storcken (1997), "Measuring Intransitivity," *Mathematical Social Science*, 34, 125–52.
- Nordgren, Loran F. and Ap Dijksterhuis (2009), "The Devil Is in the Deliberation: Thinking Too Much Reduces Preference Consistency," *Journal of Consumer Research*, forthcoming.
- Nunnally, Jum C. and Ira H. Bernstein (1994), *Psychometric Theory*, New York: McGraw-Hill.
- Paivio, Allan (1986), *Mental Representations: A Dual Coding Approach*, New York: Oxford University Press.
- Panksepp, Jaak (2004), *Affective Neuroscience*, New York: Oxford University Press.
- Peters, Ellen and Paul Slovic (2000), "The Springs of Action: Affective and Analytical Information Processing in Choice," *Personality and Social Psychological Bulletin*, 26 (12), 1465–75.
- Pham, Michel T. (2004), "The Logic of Feeling," *Journal of Consumer Psychology*, 14 (4), 360–69.
- (2007), "Emotion and Rationality: A Critical Review and Interpretation of Empirical Evidence," *Review of General Psychology*, 11, 155–78.
- Pham, Michel T., Joel B. Cohen, John W. Pracejus, and G. David Hughes (2001), "Affect Monitoring and the Primacy of Feelings in Judgment," *Journal of Consumer Research*, 28, 167–88.
- Piaget, Jean (1969), *The Psychology of the Child*, New York: Basic Books.
- Prelec, Drazen and George Loewenstein (1998), "The Red and the Black: Mental Accounting of Savings and Debt," *Marketing Science*, 17 (1), 4–28.
- Schwarz, Norbert, Herbert Bless, Fritz Strack, Gisela Klumpp, Helga Rittenauer-Schatka, and Annette Simons (1991), "Ease of Retrieval as Information: Another Look at the Availability Heuristic," *Journal of Personality and Social Psychology*, 61, 195–202.
- Shiv, Baba and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of Consumer Research*, 26 (December), 278–92.
- Shiv, Baba, George Loewenstein, Antoine Bechara, Hanna Damasio, and Antonio Damasio (2005), "Investment Behavior and the Negative Side of Emotion," *Psychological Science*, 16 (6), 435–39.
- Siemer, Matthias and Rainer Reisenzein (1998), "Effects of Mood on Evaluative Judgments: Influence of Reduced Processing Capacity and Mood Salience," *Cognition and Emotion*, 12 (6), 783–805.
- Simonson, Itamar (2008), "Will I Like a 'Medium' Pillow? Another Look at Constructed and Inherent Preferences," *Journal of Consumer Psychology*, 18 (3), 155–69.
- Slater, Patrick (1961), "Inconsistencies in a Schedule of Paired Comparisons," *Biometrika*, 48, 303–12.
- Sloman, Steven A. (1996), "The Empirical Case for Two Systems of Reasoning," *Psychological Bulletin*, 119 (1), 3–22.
- Sopher, Barry and Gary Gigliotti (1993), "Intransitive Cycles: Rational Choice or Random Error? An Answer Based on Estimation of Error Rates with Experimental Data," *Theory and Decision*, 35, 311–36.
- Stephen, Andrew and Michel T. Pham (2008)), "On Feelings as a Heuristic for Making Offers in Ultimatum Negotiations," *Psychological Science*, 19, 1051–58.

- Trope, Yaacov and Thomas Alfieri (1997), "Effortfulness and Flexibility of Dispositional Judgement Processes," *Journal of Personality and Social Psychology*, 73 (4), 662–74.
- Tversky, Amos (1969), "Intransitivity of Preferences," *Psychological Review*, 76 (1), 31–48.
- Vohs, Kathleen D., Roy F. Baumeister, and George F. Loewenstein (2007), *Do Emotions Help or Hurt Decision Making? A Hedgedoxian Perspective*, New York: Russell Sage.
- Wertenbroch, Klaus (2003), "Self-Rationing: Self-Control in Consumer Choice," in *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*, ed. George Loewenstein, Daniel Read, and Roy Baumeister, New York: Russell Sage, 491–516.
- Wilson, Timothy D., Douglas J. Lisle, Jonathan W. Schooler, Sara D. Hodges, Kristen J. Klaaren, and Suzanne J. LaFleur (1993), "Introspecting about Reasons Can Reduce Post-choice Satisfaction," *Personality and Social Psychology Bulletin*, 19, 331–39.
- Wilson, Timothy D. and Jonathan W. Schooler (1991), "Thinking Too Much: Introspection Can Reduce the Quality of Preferences and Decisions," *Journal of Personality and Social Psychology*, 60 (2), 181–92.
- Zajonc, Robert B. (1998), "Emotions," in *Handbook of Social Psychology*, Vol. 1, ed. Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, Boston: McGraw-Hill, 591–632.



How genes make up your mind: Individual biological differences and value-based decisions

Thomas Z. Ramsøy *, Martin Skov

Decision Neuroscience Research Group, Copenhagen Business School, Denmark

Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Denmark

ARTICLE INFO

Article history:

Received 16 February 2008

Received in revised form 16 November 2008

Accepted 16 December 2009

Available online 14 April 2010

JEL classification:

D01

PsycINFO classification:

2500

Keywords:

Value-based decision making

Emotions

Imaging genetics

Neuroeconomics

ABSTRACT

Neuroeconomics is the multidisciplinary study of value-based decision-making. One of the core topics is how emotions affect decision-making. Developments in economic models of decision-making have been influenced by technological innovations and empirical findings in cognitive neuroscience. Now, a recent approach in cognitive neuroscience, often referred to as “imaging genetics”, promises to make significant contributions to our understanding of both behavioral and neural aspects of value-based decision-making. Recent work has demonstrated the role of neurotransmitter alterations in clinical states such as Parkinson’s disease, depression and anxiety, and how this may affect decision behavior. However, these insights are limited through their focus on extreme neuropathology, which sheds little light on similar functions in healthy individuals. Here, we present and discuss studies of the role of drug-induced and genetically driven changes in neurotransmitter levels, and their effects on value-based decision-making. Following this, we argue that in healthy subjects, individual variance in decision behavior can be explained by such genetic factors, and gene-environment interactions. We suggest that this development should be used in neuroeconomic research in order to both improve behavioral models, by stressing the biological nature of individual variance, and through the improvement of our general understanding of the brain basis of value-based decision-making.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Neuroeconomics is the study of the neurobiology of value-based decision-making. The notion of “value-based decision-making” (VBDM) refers to a suite of functional neural processes involved in representing internal and external states of the organism, in the valuation of possible behavioral options, and in selecting motor actions based on these valuations (Rangel, Camerer, & Montague, 2008). A key assumption of most neuroeconomic models is that VBDM involves both cognitive and emotional components (Bechara, Damasio, & Damasio, 2000; Bechara, Damasio, & Damasio, 2003; Dunn, Dalgleish, & Lawrence, 2006; Fellows, 2007; Loewenstein, Weber, Hsee, & Welch, 2001). Emotions are typically seen as crucial to the generation of motivational states and the valuation of decision options, whereas cognitive processes are assumed to be involved in information processing functions such as memory, executive function and action selection. The interaction between emotional processes and cognitive processes is recognized as being complex, and it has been demonstrated that cognitive representations such as semantic knowledge about an object can modulate the emotional processes involved in the

* Corresponding author. Address: Decision Neuroscience Research Group, AØ, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark.

E-mail address: thomasr@drcmr.dk (T.Z. Ramsøy).

computation of its value (Kirk, 2008; Kirk, Skov, Hulme, Christensen, & Zeki, 2009; McClure et al., 2004; Plassmann, O'Doherty, Shiv, & Rangel, 2008). Conversely, early and unconscious affective responses in structures such as the amygdala can affect perceptual processing, and hence influence how cognitive representations are computed (Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004). In general, neuroeconomics assumes that predicted values – what is sometimes referred to as the predicted utility (Kahneman, Wakker, & Sarin, 1997) – of available options determine decision behavior. However, it has been demonstrated that this relationship between predicted value and decision can be dissociated (Weber & Johnson, 2009) and even break down in pathological situations such as addiction or pathological gambling (Berridge & Aldridge, 2008).

To date most neuroeconomic research has focused on illuminating how various factors change the workings of the emotional and cognitive processes involved in VBDM. To name just a few examples, decisions have been shown to be influenced by attentional processes, including perceptual scan strategies, assessment of the situation (framing, goal elicitation), and by the selection of information, but research has also shown them to be influenced by the physiological state of the organism (conditioning its behavioral motivation), and the hedonic values of the behavioral options (previous experience, the certainty of the outcome, the temporal availability of the reward, comparison of different possible reward possibilities, etc.). In this research, it is often assumed (implicitly or explicitly) that the neurobiological systems involved in computing these factors are uniform across the population. This is, however, not the case. The molecular make up of both emotional and cognitive processes may vary from individual to individual due to systematic biological factors. This individual variance naturally influences such processes, and have an impact on decision-making. To improve models of VBDM, and make them ecologically valid, the relevant biological factors associated with individual variance need to be included.

Here, we wish to draw attention to a number of these factors. We will present and discuss recent insights into the relationship between neurotransmitter functions, emotions and decision-making. Specifically, we focus on how genetic variation causes variations in the molecular biology of neural structures associated with VBDM. Finally, we discuss the implications of such insights on models and assumptions in neuroeconomics. Before embarking on this discussion, however, we will review how different neural structures and processes are thought to be involved in value-based decision-making.

2. Neural structures involved in value-based decision-making

Studies of patients with brain lesions provided early indications of the relationship between specific brain regions and decision-making. Since the emergence of neurology in the late 19th century, studies of brain injured patients have demonstrated a link between emotional responses and decision-making behavior. For instance, patients with lesions to the ventromedial pre-frontal cortex (vmPFC) demonstrate significant deficits in decision-making tasks including gambling (Bechara, 2003; Bechara & Damasio, 2002; Bechara, Damasio, Tranel, & Anderson, 1998; Bechara, Tranel, & Damasio, 2000b; Cavedini, Riboldi, Keller, D'Annunzi, & Bellodi, 2002; Clark et al., 2008). Other kinds of lesions, such as lesions to the amygdala, have been demonstrated to produce similar behavioral deficits (Bar-On, Tranel, Denburg, & Bechara, 2003; Bechara, 2001; Bechara & Damasio, 2002). But it was not until the late 1980s and early 1990s that it was formally suggested that decision-making was the result of an integration of cognitive and emotional processes. In particular, Damasio and colleagues proposed that patients with lesions to the vmPFC performed poorly on gambling tasks due to reduced emotional responses (Bechara, Damasio, Damasio, & Lee, 1999; Bechara, Damasio, Tranel, & Damasio, 1997). On this view, emotions provide somatic markers to guide the choosing between different available options. Bodily responses were used during the decision-making process to help labeling some choices as "bad", especially in gambles with unknown, or risky, outcomes (Bechara & Damasio, 2005; Damasio, 1996). Later studies have complicated this initial hypothesis somewhat. For example, patients with amygdala or vmPFC lesions perform better than normal controls on gambling tasks that reward risk taking (Shiv, Loewenstein, Bechara, Damasio, & Damasio, 2005), and patients with bilateral damage to vmPFC make more utilitarian moral judgments than healthy subjects (Koenigs et al., 2007). Nevertheless, the early patient studies demonstrated an advantageous role of emotion in decision-making, and have led to a general acceptance of the idea that brain regions responsible for generating emotional responses serve some fundamental role in the evaluation of stimuli and events, with the purpose of guiding or motivating decisions (Bechara, 2003).

Today, there is an immense focus on the role of emotions during decision-making processes. Emotions are often divided into positive and negative responses. However, one should note that the term "emotion" itself is unclear. Indeed, in neuroeconomics emotional responses may refer to the mental processes during expected outcomes, during the decision-making phase, or during the outcome phase. Analogous to this, in the economic literature, there has been a suggestion that one should distinguish between different forms of processes during the decision-making process, including "expected utility", "decision utility", "outcome utility," and "experienced utility" (Kahneman, 1994; Kahneman & Tversky, 1979, see also Peterson, 2007). Conversely, in cognitive neuroscience, the term "emotion" is normally used to refer to the immediate behavioral responses of an organism to stimuli or events that are, e.g., life-threatening or rewarding. Compared to the multiple utility functions provided in economics, the cognitive neuroscience approach seems far too simplistic. In defining emotions, there is indeed a need for a cross-talk between traditional disciplines. In particular, as the emerging multidisciplinary collaboration between economics, psychology and neuroscience continues, such inconsistencies need to be mapped out and resolved into a coherent and common framework.

Within the tradition of cognitive neuroscience, positive emotions are normally thought to be rooted in reward and approach behaviors, and are typically associated with neural processes taking place in the mesolimbic and mesocortical pathways (projections from the ventral tegmentum of the midbrain to the frontal lobe), processes that mainly involve neurotransmitters such as dopamine and opioid (Adcock, Thangavel, Whitfield-Gabrieli, Knutson, & Gabrieli, 2006; Adinoff, 2004; Berridge, 2007;

Pfaus, 2009; Schott et al., 2008), and more recently the neuropeptide orexin/hypocretin (Aston-Jones, Smith, Moorman, & Richardson 2009; Harris & Aston-Jones 2006). The structures implied in the dopaminergic network include the ventral striatum (VS), amygdala, orbitofrontal cortex (OfC) and anterior cingulate (ACC), and the role of this system can be tentatively described as the computing of reward prediction and of the experienced hedonic value of the actual reward outcome (Berridge, 2007; Berridge & Aldridge, 2008; Cooper & Knutson, 2008; McClure, Daw, & Montague, 2003). Several studies have uncovered the brain basis of some of the factors influencing reward based economic behavior, including risky choices, time discounting, as well as co-operation and altruism (Sanfey & Chang, 2008; Sanfey, Loewenstein, McClure, & Cohen, 2006; van 't Wout, Kahn, Sanfey & Aleman, 2006; Vorhold, 2008; Zak, 2004; Zak & Fakhar, 2006; Zak, Kurzban, & Matzner, 2004).

Negative emotions, on the other hand, are thought to be related to aversion and avoidance behaviors, and mainly driven by responses of the amygdala, lateral OfC, insula, and hypothalamus (Charney, 2003; Fischer, Andersson, Furmark, & Fredriksson, 2000; Lane et al., 1997; Price, 1999; Royet et al., 2000). The neurotransmitter serotonin is thought to mediate functions of this circuit, as alterations in the levels of serotonin has been linked to affective functions, with high levels generally being associated with increased vigilance, fear and anxiety, and low levels generally associated with reduced vigilance, depression and increased pain sensitivity (Charney, 2003; Hariri et al., 2002; Hurlemann et al., 2009; Surguladze et al., 2008; Williams et al., 2009).

A currently unresolved controversy concerns the exact role played by positive and negative emotions in decision-making processes. Many decisions seem to involve both positive and negative dimensions. For instance, choosing to buy a chocolate bar both includes the (expected and experienced) reward of eating the chocolate and the negative consequences associated with future health and obesity considerations. Similarly, gambling situations are by nature composed of positive and negative factors, such as the expected likelihood of winning or losing a bet. Gambling decisions are therefore influenced both by positive emotions such as reward expectancy and negative emotions such as aversion to loss, risk or outcome ambiguity. As a result, one may suggest that decision-making is a tug-of-war between positive and negative emotional processes. However, this idea has recently been challenged. For example, in a recent neuroimaging study Tom, Fox, Trepel, and Poldrack (2007) found that both potential rewards and potential losses were related to a scaled response in midbrain dopaminergic regions and their targets, such as the VS. Increased activation in these regions was found for higher potential gains, and reduced VS activation was found for potential losses. Surprisingly, no separate activations were specific for potential losses, as one might have expected from the models positing a separate neuroarchitecture of the aversion system. This suggests that expectation of monetary losses and gains might be fully processed by a unitary (appetitive) system, centered on the striatum. Only a few reports have shown increased activations to experienced or expected financial loss, such as the amygdala (Yacubian et al., 2006) and insula cortex (Knutson & Bossaerts, 2007; Knutson, Rick, Wimmer, Prelec, & Loewenstein, 2007; Kuhnen & Knutson, 2005). Furthermore, Clark et al. (2008) recently demonstrated that patients with lesions to vmPFC and insula displayed selective and distinctive disruptions of betting behavior. On the one hand, vmPFC damage was associated with increased betting regardless of the odds of winning, while insular damage was related to a failure to adjust bets by the odds of winning. This study clearly demonstrates the necessary role of structures implied in aversive responses, such as the insula, in decision-making under risk.

These conflicting findings may in part be explained by the lack of common definitions, in particular regarding the different kinds of utility involved in valuation and decision-making. In particular, while Tom et al. (2007) focused on decision utility, Knutson et al. (2007) focused on expected utility. Yet other studies, such as Clark et al. (2008) could not distinguish between different stages of the decision-making process. Such discrepancies need to be resolved when comparing studies, but also provides new possibilities for comparing different stages of the valuation and decision-making process.

The lack of consistent labeling and definition across disciplinary borders, is further fueled by findings that break with the traditional dichotomy between reward and aversion structures. For example, Levita et al. (2009) recently reported that the VS responded just as strongly – or even stronger – to the expectation of punishment or pain, as to reward expectation. Furthermore, studies have suggested that dopamine and the VS is more involved with outcome expectancy rather than purely with liking and hedonic pleasure (O'Doherty et al., 2002; Suri, 2002), and that the opioid system and structures such as the OfC may be involved in the experience of hedonic reward (O'Doherty et al., 2002; Peciña, Smith, & Berridge, 2006; Petrovic et al., 2008). Similarly, findings imply that the amygdala is also involved in generating positive emotions (Murray, 2007). Indeed, these studies demonstrate that structures such as the amygdala and VS in some situations demonstrate a non-linear U-shape response function, with high activation levels for positive and negative emotions, and low activation for neutral events. While the nature of such findings are yet poorly understood, it is possible that these findings may in part be explained by complex excitatory and inhibitory interconnections between the amygdala and VS. For now, the neuroimaging studies cited here imply that there is as of yet no clear distinction in the neural architecture between positive and negative emotions. Furthermore, the distinctions between different forms of utility, as used in economics, may be informative. It is possible that structures such as the amygdala may be selectively important for some stages of the VBDM process. Whether this is the case, or whether these structure-function relationships are more general in nature, remains to be seen.

The debate between “separatists” and “unitarists” may be informed by on-going developments in neuropharmacology and in the combination of cognitive neuroscience, neuroimaging and genetics – often referred to as imaging genetics (Hariri, Drabant, & Weinberger, 2006; Meyer-Lindenberg & Weinberger, 2006). In the rest of this paper we present and discuss studies that relate neurotransmitters to gambling behavior, by first introducing the link between neurotransmitter levels and pathological states of emotion and behavior. Furthermore, we will look at how artificially induced changes in levels of dopamine and serotonin may affect emotion and decision-making. We then present studies of the genetic influence on individual

differences in dopamine and serotonin levels, and how this has been found to affect emotions and decisions. Intriguingly, as we will show, changes in levels of neurotransmitters such as serotonin and dopamine have different effects on approach and avoidance behavior.

3. Neurotransmitter pathology and decision-making

The role of neurotransmitters in VBDM is clearly demonstrated by studies of pathological conditions. Although the link between neurotransmitter levels and behavior has been debated for decades, studies of mental conditions have consistently linked high and low levels of neurotransmitters such as dopamine and serotonin, to changes in mental states and behaviors. For instance, as dopamine levels decrease, due to pathological neurodegeneration of dopamine producing regions such as the brainstem substantia nigra, a person will develop symptoms of Parkinson's disease (PD). Beside progressive loss of voluntary and involuntary muscle control, PD patients demonstrate a number of secondary symptoms, such as depression and anxiety. In particular, PD is associated with marked frontal lobe impairment which gives rise to behavioral problems associated with anticipation, planning, and the regulating and directing of purposeful behavior (Azuma, Cruz, Bayles, Tomoeda, & Montgomery, 2003). Treating PD by pharmacologically increasing dopamine levels has – apart from improving the primary symptoms – also been related to significant alterations in decision-making behaviors in some patients, in particular the increased likelihood of developing pathological gambling and hyper-sexual behaviors (Cabrini et al., 2009; Merims & Giladi, 2008; Stamey & Jankovic, 2008). Thus, there seems to be a link between increased dopamine levels and decision-making behaviors and executive functions, including pathological gambling. The nature of this link to gambling behavior could be the result of increased risk taking, or reduced responsiveness to losses. Studies of pathological gambling in dopamine agonist treatment in PD show an increase in reward driven behaviors and diminished impulse control, rather than changes in responses to losses or aversion in general (Cilia et al., 2008; Evans, Strafella, Weintraub, & Stacy, 2009; Ferrara & Stacy, 2008; Ondo & Lai, 2008; Steeves et al., 2009; Torta & Castelli, 2008).

According to the suggested link between dopamine levels and risk taking, pathological states leading to abnormally high levels of dopamine should demonstrate a similar pattern. Schizophrenia is a mental disorder with increased dopamine levels, or at least increased dopaminergic transmission in some brain regions (Meyer-Lindenberg et al., 2002), that is characterized by distorted perception and thoughts, and abnormalities in the understanding and expression of reality (Antonova, Sharma, Morris, & Kumari, 2004; Frith & Dolan, 1996; Gallhofer, Lis, Meyer-Lindenberg, & Krieger, 1999). Recently, researchers have demonstrated decision-making impairments in schizophrenia (Desai & Potenza, 2009; Yip, Sacco, George, & Potenza, 2009). In particular, it has been suggested that schizophrenics display an impaired ability to integrate affective and cognitive information (Heerey, Bell-Warren, & Gold, 2008) and that schizophrenic patients display many of the same characteristics as pathological gamblers (Borras & Huguelet, 2007).

Pharmacological adjustments of dopamine levels are known to affect psychiatric and neurological symptoms: one of the side effects of dopamine antagonism treatment of schizophrenia is Parkinsonism, while side effects of increasing dopamine levels in PD include psychosis. In addition, behavioral symptoms of affective responses and decision-making seem to follow a similar scaling effect with dopamine levels. These results suggest that increased levels of dopamine affect decision-making behavior through the uncoupling of affective and cognitive elements. In healthy subjects, decision-making has been found to be based on the integration of emotional and cognitive components. This link is thought to be represented in the ventromedial prefrontal cortex, as lesions to this region affect normal decision-making (Bechara, 2001; Bechara, 2003; Bechara, Tranel, & Damasio, 2000a; Clark et al., 2008). Conversely, studies of low levels of dopamine have been less conclusive. Untreated PD patients (with low dopamine levels) show increased anxiety or depression, and one may expect that this will affect their decision-making, i.e., by making them more risk averse or generally more responsive to aversive outcomes. Also, pharmacological adjustments of dopamine levels to improve PD and schizophrenia seem to suggest that dopamine is a scaling variable for decision-making behavior. The clinical observation that a side effect of dopamine antagonism in schizophrenia is parkinsonism (both primary and secondary symptoms), while side effects of increasing dopamine levels in PD include psychosis, pathological gambling and hyper-sexuality, seems to support this notion.

Brain regions involved in aversion have been shown to be modulated by the neurotransmitter serotonin (Kalbitzer et al., 2009; Watson, Ghodasra, & Platt, 2009; Zuckerman, Ballenger, & Post, 1984). The role of serotonin in emotional processing has been demonstrated in a range of studies during the past several decades. All have consistently related changes in serotonin levels to emotional disturbances such as anxiety, depression, obsessive-compulsive disorder and impulse control disorders (e.g., de Win et al., 2004; Kalbitzer et al., 2009; Merens, Willem Van der Does, & Spinhoven, 2007; Nugent et al., 2008; Reimold et al., 2008; Riedel, Klaassen, & Schmitt, 2002; Ruhé, Mason, & Schene, 2007; Savitz, Lucki, & Drevets, 2009; Young & Leyton, 2002). Drugs that increase the amount of synaptic serotonin, such as the selective serotonin re-uptake inhibitor (SSRI), lead to reduced depression and anxiety (Gorman & Kent, 1999; Sokolski, Conney, Brown, & DeMet, 2004; Spillmann et al., 2001). Such treatments also improve performance on decision-making tasks, possibly because decisions are now less driven by over-activation of the aversion circuit (Dannon, Lowengrub, Gonopolski, Musin, & Kotler, 2005; Figgitt & McClellan, 2000; Hollander et al., 2000). This suggests that serotonergic transmission influences emotional processing and decision-making, and that it may serve as an important factor influencing VBDM.

Altered levels of serotonin have thus been linked to mood disorders such as depression and anxiety, although this link between disorder and neurotransmitters may be too simplistic. For example, depressed patients demonstrate both global

and regional decreases in brain serotonin levels (Meyer et al., 2001; Meyer et al., 2004). Depression is a mental disorder well-known for being characterized by a low mood, low self-esteem, and a loss of interest or pleasure in normally enjoyable activities. A fundamental aspect of depression is a general reduced sensitivity and response to rewards (Foti & Hajcak, 2009). Given these findings, one may assume that depressed subjects will display altered decision-making behavior. Such a relationship has indeed been found. Studies have reported reduced sensitivity to rewards and biased expectation to negative outcome in depressed subjects (Foti & Hajcak, 2009). Consistent with this view, yet perhaps surprisingly, depressive subjects perform better on the Iowa gambling task (IGT) (Smoski et al., 2008). In this card drawing task, choices with high payoff are also associated with rare but significant losses, with a cumulated net loss. Conversely, alternate options are associated with relatively lower immediate payoff but also lower losses, leading to a cumulated net win. Depressed subjects were faster at learning to avoid risky responses (with net losses) than healthy subjects, and were even found to earn more money than healthy subjects. As noted, depressed subjects respond less to rewards. This could mean that depressed subjects, compared to healthy subjects, are less responsive to reward size, and thus do not fall prey to a "gambler's fallacy" effect of going for large yet rare rewards (Burns & Corpus, 2004; Erev, 1998).

Another mood disorder influenced by serotonin, anxiety disorder, is defined as an unpleasant emotional state, for which the cause is either not readily identified, or perceived to be uncontrollable or unavoidable. Anxiety is a heterogenous disease, but can in general be seen as a future oriented cognitive and emotional state/trait characteristic involving several components, such as anxious apprehension, worry, affective or behavioral conflicts, and altered approach/avoidance behaviors. Social anxiety, the experience of emotional discomfort regarding social situations, has been related to global and regional decreases in brain serotonin levels (5-HT_{1A}) (Lanzenberger et al., 2007), and has also been demonstrated for panic disorder, characterized by recurring severe panic attacks (Nash et al., 2008). However, the exact role of serotonin in anxiety overall is still unclear. Patients with anxiety demonstrate increased vigilance and fear responsiveness, and studies have demonstrated that anxious subjects also demonstrate increased amygdala responses to aversive stimuli (Birbaumer et al., 1998; De Bellis et al., 2000; Rauch, Shin, & Wright, 2003; Stein, Goldin, Sareen, Zorrilla, & Brown, 2002). Taken together, such results suggest a link between serotonin and aversive responses, but are as of yet inconclusive. Can anxiety influence decision-making in situations that include factors such as risk, ambiguity or loss? Indeed, this has been demonstrated in recent studies of decision-making in anxious subjects. For example, in a neuroimaging study, subjects with high trait anxiety performed equally well compared with subjects with normal trait anxiety on a decision-making task with either a high or low error rate. However, during the low error rate condition, when the likelihood of incorrect responding was very low, high trait anxiety subjects demonstrated higher activation of the anterior cingulate and medial PFC. This suggests that anxious subjects devote more time and processing resources to decision-making at times when chances of errors are very low (Paulus, Feinstein, Simmons, & Stein 2004). Possibly, anxiety influences VBDM through both the regulation of general mood and vigilance as well as an increased attention to errors, and aversive responses.

When taken together, the studies discussed above suggest a role for serotonin in VBDM. Increased levels of serotonin may be related to increased anxiety, leading decisions to be influenced by fear of losing, risk or even ambiguity. Conversely, low levels of serotonin are related to depression, in which responses to rewards and losses are generally lower, thus leading decisions to be generally less influenced by emotional factors. Besides psychiatric disease, these findings hint that induced changes in serotonin levels may also influence value-based decision-making in healthy subjects. This will be demonstrated in the next section, where we look at pharmacological interventions that influence neurotransmitter levels.

4. Drug altered decisions

Value-based decisions are affected by induced alterations of neurotransmitter levels in healthy subjects. Here, two levels of action are found. *Active effects* are the immediate effects of the drug, with an effect normally lasting for minutes to hours. *Long-term effects* are the changes to the neurotransmitter system after prolonged and/or repeated exposure to a drug, with a time span of weeks to months and years.

As described previously, one side effect of dopamine agonist treatment in Parkinson's disease include pathological gambling. Similarly, increases in dopamine levels in other subject groups have also been reported to alter decision-making behavior. For example, the psychoactive drug cocaine is known to block dopamine re-uptake, leading to a general increase in synaptic dopamine levels. Active effects of cocaine have been shown to affect decision-making behavior, including increased risk taking during game playing (Hall et al., 2000; Monterosso, Ehrman, Napier, O'Brien, & Childress, 2001; Stalnaker et al., 2007), and reckless behaviors (MacDonald et al., 2008). Furthermore, in a study by Sevy et al. (2006), healthy subjects were given a mixture containing branched-chain amino acids (BCAA), valine, isoleucine and leucine, or a placebo treatment. The BCAA treatment was found to decrease dopamine levels, and to impair performance on the IGT. Specifically, the authors suggest that the BCAA treatment impaired IGT performance through an increased attention to short-term outcomes, at the cost of more long-term learning of outcome contingencies. As the IGT is a decision-making test that relies on normal arousal functions, reduced dopamine levels may lead to impaired emotional decision-making characterized by shortsightedness and impulsivity. Contrary to this, based on the previously noted IGT behavior in depressed subjects, it is possible that the observed behavioral changes following dopamine antagonism can be explained by an increased sensitivity to losses. However, as depressed subjects demonstrated an actual improvement in IGT performance, these results are not directly comparable.

In the same way, just as models of the role of serotonin may focus on chronic levels of serotonin, as seen in studies of mood disorders, it is also possible to study the effects of acute and induced changes in serotonin function. Studies of serotonin depletion have largely confirmed the studies of VBDM in anxiety and depression. For example, tryptophan depletion (TD) is a non-medical method for lowering the levels of serotonin through a diet with all amino acids except tryptophan. Studies using positron emission tomography (PET) have shown that TD leads to significant decrease in brain 5-HT₂ receptor binding in various cortical regions including the insula (Yatham et al., 2001), but no changes in regional serotonin transporter binding potential (5-HTT BP), an index of 5-HT density and affinity decreased plasma free tryptophan (Praschak-Rieder et al., 2005). Behaviorally, while leaving subjective mood ratings relatively intact, tryptophan depletion may alter emotional processing such as decreased responses to aversive stimuli, including reduced fear and anxiety responses (Harmer, Rogers, Tunbridge, Cowen, & Goodwin, 2003; Kähkänen et al., 2005; Munafò, Hayward, & Harmer, 2006; Roiser, Müller, Clark, & Sahakian, 2007). Given the role of the aversion circuit in the mediation of such responses and behaviors, one would expect the effect of tryptophan depletion to be involved in changes in this network. Indeed, Cools and colleagues (2005b) found that tryptophan depleted subjects showed stronger neural responses in traditional aversion structures such as the amygdala, and that these responses were related to self-reported threat sensitivity. Similar findings have been reported for other reinforcement learning paradigms (Evers et al., 2005; Finger et al., 2007; Rogers et al., 2003).

A potent model of chronic serotonin depletion can be found in subjects using recreational drugs that, as a long-term consequence, reduce serotonin levels. Such serotonin level reductions occur through long-term exposure to serotonin agonists. Recreational drugs like 3,4-methylenedioxymethamphetamine (MDMA, commonly known as ecstasy) and hallucinogens (Hal) exert their main psychoactive effects through modulations of the serotonin (5-HT) neurotransmitter system, in particular via the 5HT_{2A} receptor (Montoya, Sorrentino, Lukas, & Price, 2002; Morton, 2005; Parrott, 2001). Recently, we have reported a general *in vivo* decrease in 5-HT_{2A} in MDMA/Hal users compared to non-users, where the strongest effects were found in the prefrontal cortex, medial temporal lobe (in particular hippocampus and amygdala) and putamen (Erritzøe et al., 2008). Studies of changes in the neural architecture following MDMA/Hal recreational drug use suggest large effects on regions involved in emotional processing. In a comparison of MDMA/Hal recreational users and non-users, we have recently reported (Ramsøy et al., 2009b) that MDMA/Hal users demonstrate significant grey matter volume reductions in the OfC and ACC, as well as reduced white matter the uncinate fasciculus, a structure implicated in emotional processing. In particular, we have also found that MDMA/Hal users show altered emotional responses, as reduced neural responses to aversive faces in the left amygdala and left rhinal cortex, and an increased response in the left medial OfC (Ramsøy et al., 2009a). MDMA/Hal users demonstrate altered decision-making behavior, in particular through a generally elevated behavioral impulsivity, leading to a poorer performance on the IGT (Quednow et al., 2007). This supports the link between serotonin levels and VBDM, in that lower levels of serotonin may lead to increased impulsivity, and higher levels may lead to decisions being more driven by fears and aversions.

5. Genes and rewards

In very much the same way that both pathological states and induced neurotransmitter level changes can influence brain function and behavior, there are naturally occurring differences in neurotransmitter levels. As mentioned, imaging genetics is the study of how genetic differences lead to individual differences in the morphology and functions of the brain, and differences in behavior (see, as a first introduction, Canli et al., 2005; Hariri & Holmes, 2006; Hariri et al., 2006; Winterer, Hariri, Goldman, & Weinberger, 2005). Imaging genetics research includes structural and functional brain imaging tools as phenotypic assays to evaluate genetic variation. Put differently, the focus is on how genetic variations, in combination with life experience, lead to measurable changes in cognitive and affective processes, and eventually behavior (including decision-making). In practice, studies in imaging genetics combine genetic analysis, behavioral data and neuroimaging to assess how genetic differences influence brain function and behavior.

As a principle, genes influence behavior indirectly, through the modulation of the molecular environment of neurons. The coding sequence of a gene leads to the synthesis or function of proteins, and this affects the operation of a cell. Such a protein may, for example, function as a membrane gateway for specific molecules, or may act as a transporter of synaptic neurotransmitters. In this way, genes can code for parts that play specific functions in a neuron. Consequently, individual differences in specific genes may lead to measurable differences in protein synthesis and, through this, to differences in how neurons respond. At larger scales, such changes might lead to differences in brain activations and, ultimately, overt behavior. Genetic variations – so-called polymorphisms – can be situated in genes encoding different proteins important for the neurotransmitter system, such as proteins involved in the synthesis, transport, post-synaptic uptake, pre-synaptic re-uptake and breakdown of the neurotransmitter. As different genes regulate the molecular biology of these processes, individual differences occur at any level of the neurotransmitter process.

As described above, studies have linked increased dopamine levels and risk behavior in gambling, and several other studies have demonstrated a close relationship between dopamine levels and impulsivity (Cardinal, Winstanley, Robbins, & Everitt, 2004; Pattij & Vanderschuren, 2008; Takahashi, 2007; Volkow, Fowler, Wang, Baler, & Telang, 2009; Volkow, Fowler, Wang, Swanson, & Telang, 2007). This link would suggest that genetic variance in dopamine levels would influence risk willingness during gambles. Indeed, several studies now suggest this to be the case. Catechol-O-methyl transferase (COMT) is an enzyme that degrades catecholaminergic neurotransmitters such as dopamine, epinephrine, and norepinephrine

(Meyer-Lindenberg et al., 2005; Meyer-Lindenberg et al., 2006; Strauss et al., 2004; Volavka, Bilder, & Nolan, 2004). This means that a higher operation level of COMT would lead to greater breakdown rates of dopamine, and thus generally lower levels of synaptic dopamine.

A functional polymorphism in the COMT gene, the val¹⁵⁸met variation, is the cause of a fourfold increase in enzyme activity (Tan et al., 2007). In particular, the low-activity met allele has been associated with improved working memory performance (Goldberg et al., 2003), while some val allele carriers display an advantageous increase in emotional resilience against anxiety and dysphoric mood (Enoch, White, Harris, Rohrbaugh, & Goldman, 2002; Zubieta et al., 2003). In a recent study, researchers tested subjects with different allelic COMT variations during an inter-temporal choice paradigm (Boettiger et al., 2007). By comparing subjects according to their genotype, the researchers found a positive correlation between the choice of immediate reward and the magnitude of BOLD (blood oxygen level dependent) fMRI (functional magnetic resonance imaging) signal during decision-making in the posterior parietal cortex (PPC), dorsal PFC, and rostral parahippocampal cortex, while the tendency to choose larger, delayed rewards was correlated with activation in lateral OFC. By dividing their subjects into their allelic types of the COMT gene, the researchers were able to look at differences in brain activation and behavior. It was found that the genotype at the val¹⁵⁸met polymorphism of the COMT gene predicted both impulsive behavior and activity levels in the dorsal PFC and PPC during decision-making. In particular, this difference was driven by the homozygote val/val genotype compared to homozygote met/met and heterozygote val/met genotypes. That is, val/val subjects more often displayed increased levels of activity in dorsal PFC and PPC, and more often chose immediate rewards. In this sense, individual differences in choosing immediate or delayed rewards are shaped by each individual's genetic make up.

Recent evidence suggests that a synonymous polymorphism within the COMT gene (rs4818 C/G) accounts for a greater variation of COMT activity than the functional val¹⁵⁸met polymorphism (Diatchenko et al., 2005). In a study of the C/G rs4818 variance, Roussos, Giakoumaki, Pavlakis, and Bitsios (2008) found that the C/G polymorphism produced strong and differential effects on PFC functions and behavior. The G/G allelic variance is associated with the highest levels of COMT activation, a >18-fold increase, and thus the lowest dopamine levels compared to the C/G and C/C types. By comparing subjects on two problem solving tasks – one relying on high arousal (the IGT) and one low arousal effects (Stockings of Cambridge) – the researchers found a clear dissociation between genotype on the two tasks. On the Stocking of Cambridge task, C/C individuals demonstrated the best performance, G/G the worst and with the C/G individuals as intermediate. However, during the IGT, G/G individuals performed best, while C/C individuals performed the worst and C/G at intermediate levels. This suggests that prefrontal dopamine levels influence task performance differently, and that low levels are disadvantageous for planning in non-arousing problem solving but optimal in arousal based decision-making. Contrary, high prefrontal dopamine levels may be advantageous for non-arousing problem solving, but disadvantageous for problems with high arousal.

To sum up, genetic differences in the COMT system produce differences in the availability of dopamine in the brain, and through this variance influences decision-making. Other genetic polymorphisms influence the dopaminergic system, and have been identified as likely contributors to differences in decision-making aspects such as impulsivity (Kreek, Nielsen, Butelman, & LaForge, 2005). Hence, the reward seeking aspect of decision-making under risk (such as gambles) seems to be influenced by dopamine levels. High levels of dopamine may be related to more risky gambles, and poor performance on non-emotional tasks, but may show benefits on other tasks that involve arousal, such as the IGT.

6. Genes and aversion

The serotonergic system is also influenced by individual genetic differences at all stages (Hariri et al., 2006). For example, humans show a common variation of a promoter region of the serotonin transporter gene called 5-HTTLPR: a short (s) and a long (l) version. People who have two copies of the long genetic sequence (homozygote l/l) in this region are thought to have less synaptic serotonin, due to a higher pre-synaptic re-uptake of the neurotransmitter. Conversely, people with two copies of the s allele (homozygote s/s) show a lower re-uptake and therefore have higher synaptic serotonin levels. Thus, emotional responses in a population may show individual variance according to 5-HTTLPR types. More specifically, the homozygote s/s allele version is associated with increased activations of structures innervated by the serotonergic system, including the amygdala, lateral OfC, insula, and hypothalamus.

In an fMRI study Hariri et al. (2002) compared the activation of the amygdala in the 5-HTTLPR allelic groups during an emotion task (matching of facial expressions compared to matching of geometric figures). This task is well known to produce a robust activation of the amygdala (LeDoux, 1993; LeDoux, 2003; Vuilleumier et al., 2002). By comparing the aversion response in s/s and l/l subjects, the researchers found that amygdala activation differed according to genotype. The s/s group showed a significantly higher signal increase in the amygdala during the emotion task than the l/l group. In other words, the level of amygdala activation was influenced by genotype, through a variation in the availability of synaptic serotonin. The results imply that the higher concentrations of serotonin found in the s/s group lead to higher engagement of the amygdala, and that this elevated activation produces a heightened emotional response to aversive stimuli. In behavioral terms, it may be contended that subjects differ in the decision-making process to aversive stimuli – e.g., aversion for loss or risk – in a way that is influenced by their genetic make up.

The question thus arises whether genetic polymorphism of the serotonin system influences decision-making behaviors. Following the studies of how mood disorders, pharmacological challenges and genetic differences affects emotional

responses, one would certainly expect that genetically induced differences in serotonin levels may lead to measurable changes in VBDM. As demonstrated previously, studies of pharmacological challenges provide strong indications that a major role of serotonin is in determining sensitivity to reinforcement information (Cools et al., 2005a; Evers et al., 2005; Finger et al., 2007; Rogers et al., 2003) including social reinforcement information (Harmer et al., 2003; Marsh et al., 2006). Thus, it would be expected that similar effects should be found for differences in serotonin function that are genetically mediated. In a study of the effects of the 5-HTTLPR genotype, it was found that genotype did indeed modulate long-term decision-making in humans and rodent, including complete and heterozygous knock-out models (Homberg, van den Bos, den Heijer, Suer, & Cuppen, 2008). Using the IGT, the researchers found that subjects homozygous for the short allele of the 5-HTTLPR chose more disadvantageously than subjects homozygous for the long allele. In rodents, similar trends were found in a comparable decision-making task, suggesting that genetically mediated levels of SERT impact decision-making. In particular, the researchers found that the genotype effects could be explained by differences in the responsiveness to outcomes, and how these ultimately guided subsequent choices. This suggests that decisions altered through SERT manipulation occur through changes in emotional responses as well as top-down modulation of emotional responses, such as PFC-amygdala functional connectivity. Indeed, this was recently observed in an fMRI study comparing subjects homozygous for the short vs. long allele of the 5-HTTLPR gene (Roiser et al., 2009), in which subjects were tested on a gambling task that was either framed in terms of gains or losses. Short allele homozygotes demonstrated a stronger framing effect, and demonstrated a higher amygdala response while making choices in accord with the frame. Furthermore, while long allele homozygotes showed a stronger PFC-amygdala coupling for choices countering the frame, no such effect was found in the short allele homozygotes. This suggests that genetic variation in the 5-HTTLPR is responsible for differences in the effect of emotion and emotion regulation on value-based decisions. It should also be noted that this finding, demonstrating a role for the aversion circuit in decision-making, conflicts with the recent claim that VBDM does not involve primary emotional regions such as the amygdala (Tom et al., 2007). However, these studies differ with respect to the aforementioned distinctions between different kinds of utilities in the decision-making process. While the IGT typically conflates many utility forms such as expected utility, decision utility and experienced utility, the Tom et al. study focused particularly on decision utility. Thus, conflicting results may only be so in a superficial way and as a result of differences of measures.

Building on the findings from pharmacological challenges, Blair et al. (2008) studied the effects of tryptophan depletion on subjects with allelic variations in the 5-HTTLPR genotype. Studies of the effects of tryptophan depletion have reported considerable variation in individual responses (Booij, Van der Does, & Riedel, 2003), and a genetic factor in the mental and behavioral responses to tryptophan depletion has been reported earlier (Neumeister et al., 2002). It was therefore assumed that a variation in responsiveness to tryptophan depletion on a decision-making task would be influenced by 5-HTTLPR genotype. Indeed, tryptophan depletion had a disproportionate effect on the performance in subjects homozygous for the long allelic version, through modulating the response to punishment information. Here, subjects with the long allelic genotype demonstrated significant difficulties in their ability to learn to avoid responding to stimuli that was associated with punishment. In addition to improving our understanding of individual variance in VBDM, and the individual effects of serotonin alteration, this study also questions the claim that such decisions are not related to emotional structures such as the amygdala (Tom et al., 2007).

7. How genes make up your mind

Understanding the brain basis of value-based decisions has the potential to improve our ability to grasp, and model, the underlying mechanisms of our behaviors. The current view on the relationship between neurotransmitters and decision-making has been heavily influenced by studies of pathology and how large changes in neurotransmitter levels (e.g. as seen in drug abuse) can make decision-making behavior go awry. Such effects may influence our general view of neurotransmitters in VBDM and our understanding of psychopathology, but shed little light on behavior in healthy individuals. Contrary to this, imaging genetics may provide an improvement in our understanding of VBDM in healthy individuals. In particular, imaging genetics studies suggest that individual differences are not randomly distributed, but may follow systematic and qualitative differences in how decisions are made. Before such progress can be made, there is a need for disentangling the different ways in which VBDM are studied and discussed. In particular, the debate on whether positive and negative emotions are separately organized in the brain, and whether they make different contributions to value-based decision-making, need to distinguish between different stages of the decision-making process. Here, there is a need to determine whether the effects of neurotransmitter variations affect value-based decisions globally, or during specific phases of the process. Both scenarios are equally likely: in the case of tryptophan depletion, it is possible that lower serotonin levels affect emotional responses equally during computations of expected utility, decision utility and experienced utility stages. Alternatively, serotonin depletion may affect selected stages of the process, such as experienced utility, without affecting other stages. There is also a different take on the distinction between global and specific effects. In MDMA/Hal users, we have reported both a global decrease in serotonin binding, as well as regional differences in such effects, such as the amygdala. The relative effects of local and global neural effects on emotional responses and decision-making are still unclear. Consequently, this warrants further studies.

The insights that imaging genetics affects emotions, value generation and decision-making have two primary effects. First, imaging genetics studies suggest that individual differences are not random or continuous, but systematic and

discontinuous. This further indicates that an agent – in economics often referred to as a representative agent – cannot be typified or standardized. Thus, while many economic models apply a standardized agent as a model, results from imaging genetics suggest that this view is unwarranted. Instead, a view encompassing individual differences should be applied. This is not to suggest that neuroeconomic models should consider all possible genetic factors and their behavioral effects. This will probably make such models far too complex and particular. Rather, genetic variance may suggest that economic models should contribute to the identification and interpretation of endophenotypes of VBDM. The term endophenotype originally stems from biology (Gottesman & Gould, 2003), and can be understood as the search to elucidate genetic associations with phenotypes (behavioral types) of interest. In this paper we have provided examples of such endophenotypes, through demonstrating how genetic effects in serotonergic and dopaminergic systems can lead to differences in neural processes and behavioral outcomes.

Furthermore, as demonstrated in studies on the difference in the COMT gene, the rs4818 C/G polymorphism accounts for significantly more than the val¹⁵⁸met polymorphism in the effect of emotion in decision-making. Through the discovery and comparison of single gene effects, neuroeconomic research may gain insights into factors that make significant and systematic impact on behavior, and ultimately improve neuroeconomic models of VBDM.

A second main impact of imaging genetics is an improved understanding of the underlying neurobiological mechanisms in VBDM. By gaining insights into the factors regulating neural functions, regional brain responses and how they affect behavior, our general understanding of the brain basis of VBDM will improve significantly. Such insights holds the promise both to improve our theoretical understanding of the mechanisms underlying decision behaviors, and to influence how we assess and treat disorders of decision-making.

In terms of the debate on the nature of the brain's valuation system, data from studies investigating the pharmacological and genetic influence on reward processing and decision-making behavior may help us to rethink the models based on neuroimaging and patient studies. These models have tended to focus on large-scale anatomical structures due to the nature of their data. Hence, the "separatists" have marshaled evidence of how certain structures appear to be implicated in approach behavior (ventral striatum and medial OfC, for instance) while others seem to subserve avoidance behavior (e.g., amygdala and insula). In contrast, the "unitarists" focus on evidence suggesting that activity in these structures underlie both approach and avoidance behavior. However, anatomical structures contain a number of distinguishable cell groups associated with different functions, and an increased understanding of their molecular nature – as evidenced by pharmacological or genetic modulations – can help illuminate this more complicated picture. For instance, in animal models it is possible to show how individual cell groups respond to neurotransmitter manipulation through the use of microinjections of agonists or antagonists. For instance, Berridge and colleagues have shown that injection of the mu-opioid agonist DAMGO into the nucleus accumbens (in the VS), leading to an increase in liking reactions to sucrose, sensitizes the cells in the rostral dorsal quadrant of its medial shell, but not cells in other parts of this structure (Peciña & Berridge, 2005). Such results suggest that the molecular "action" going on in the structures presently identified as part of the brain's valuation system is much more complex than presently understood. In order to tease out its precise computational details we need to include molecular manipulations in future decision-making studies.

8. Potential research questions

There are still many unanswered questions in the role of gene-driven variations in neurotransmitter functions, and their impact on emotions and value-based decisions. In particular, we suggest a few specific research questions that should be explored in the near future.

8.1. What is the interplay between genetic and experience based factors in VBDM behavior?

The relative influence of heritage and learning – also called the "nature–nurture debate" – has been a conundrum for anyone with a keen interest in the relationship between brain, mind and behavior. Historically, the debate has certainly been influenced by the dichotomization between inborn and learned factors: is our behavior determined by our genetic make up, or are we solely determined by our personal learning history? As such dichotomies have softened up during the past decade, and are replaced by a realization that there is no such mutually exclusion, many new and interesting discoveries have been made. More recent studies have now uncovered the relative role of genes and learning in areas such as intelligence (e.g. van der Sluis, Willemsen, de Geus, Boomsma, & Posthuma 2008) and personality (e.g. Meyer-Lindenberg et al. 2006). For example, in the study of personality development, Caspi et al. (2002) published a highly influential article on gene–environment interaction in personality development. Here, they reported that maltreated children would differ in the development of antisocial personality and violent behavior depending upon whether their genotype conferred high or low levels of MAOA expression, a neurotransmitter-metabolizing enzyme. Thus, Caspi et al. showed that a genetic variation may moderate the influence of environmental factors on behavior in a rather dramatic manner. In the study, children with a low-level MAOA genotype only developed an antisocial personality if maltreated (if a child was not maltreated, a low-level MAOA polymorphism did not cause antisocial behavior). At the same time, maltreatment did not affect children with a high-level MAOA polymorphism. Thus, neither genotype nor maltreatment were sufficient causes of antisocial behavior. Only the interaction between genetic and environmental factors was sufficient to produce antisocial behaviors.

Following this approach, there is a need for studies in VBDM research that focus on the role of genes, environment and gene–environment interaction on emotions, valuation and decision-making processes. In other words, what are the gene–environment interactions that lead to differences in valuation processes and decision-making behavior, such as whether a person becomes risk averse or risk seeking, or how pathological gambling develops?

8.2. What is the role of gene–gene interactions on VBDM behavior?

As hinted throughout in this text, there are numerous genetic influences on the function of a single neurotransmitter. As noted earlier, variations in the COMT gene (rs4818 C/G) accounts for a greater variation of COMT activity than the functional val¹⁵⁸met polymorphism, and is thus expected to have a stronger impact on COMT-related behavioral effects than the val¹⁵⁸-met polymorphism. However, one can also assume that there may be additive or subtractive effects of having one particular rs4818 C/G variant, combined with one particular val¹⁵⁸met variant. Furthermore, the interactive genetic effects on serotonin and dopamine functions (and other neurotransmitters) need to be explored with respect to VBDM behaviors. Indeed, there is an emerging literature on these interactions (De Luca, Tharmalingam, Sicard, & Kennedy, 2005; Doornbos et al., 2009; Tang et al., 2009), and such approaches hold the promise of providing further understanding of the relative role of genetic effects on emotions, valuation and decision-making. Following this strategy, studying gene–gene interaction may provide the means to better understand the relative roles of neurotransmitter levels on VBDM functions.

8.3. Gene/neurotransmitter effects and consciousness?

Decision-making is most often thought of as a conscious and explicit process, but several studies have demonstrated that consciousness during decision-making is limited, and that unconscious processes may have a significant behavioral impact (Rey, Goldstein, & Perruchet, 2009; Soon, Brass, Heinze, & Haynes, 2008; Suhler & Churchland, 2009). Differences in neurotransmitter levels, either stemming from genetic or induced effects, may operate on either a conscious or an unconscious level. Put differently, increasing the level of serotonin may make a person more loss averse at the behavioral level without the person noticing this change on a subjective level. Conversely, as dopamine has been tied to the expectation phase of value-based decision-making, it is possible that natural or induced alterations in dopamine levels may influence subjective expectations without changing overall decision behavior. As only recent studies have started to explore the relationship between value-based decision-making and consciousness, we still have a poor understanding of to what extent consciousness plays in the observed behavioral effects. In addition to the inclusion of sentience as a factor in the study of genes, neurotransmitters and VBDM, one should also bear in mind the different distinctions between different stages and forms of utility in the decision-making process. Thus, it is possible that such a research program may uncover valuable dissociations between behavioral and subjective effects.

Acknowledgement

Thomas Zoëga Ramsøy is supported by a research Grant from the National Research Council on pathological gambling (the Decision Neuroscience Project).

The authors are grateful for valuable comments and suggestions from Dr. Susanne Henningsson, and two anonymous reviewers.

References

- Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., & Gabrieli, J. D. (2006). Reward-motivated learning: Mesolimbic activation precedes memory formation. *Neuron*, 50(3), 507–517.
- Adinoff, B. (2004). Neurobiologic processes in drug reward and addiction. *Harvard Review of Psychiatry*, 12(6), 305–320.
- Antonova, E., Sharma, T., Morris, R., & Kumari, V. (2004). The relationship between brain structure and neurocognition in schizophrenia: A selective review. *Schizophrenia Research*, 70(2–3), 117–145.
- Aston-Jones, G., Smith, R. J., Moorman, D. E., & Richardson, K. A. (2009). Role of lateral hypothalamic orexin neurons in reward processing and addiction. *Neuropharmacology*, 56(Suppl. 1), 112–121.
- Azuma, T., Cruz, R. F., Bayles, K. A., Tomoeda, C. K., & Montgomery, E. B. (2003). A longitudinal study of neuropsychological change in individuals with Parkinson's disease. *International Journal of Geriatric Psychiatry*, 18(11), 1043–1049.
- Bar-On, R., Tranel, D., Denburg, N. L., & Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain*, 126(Pt. 8), 1790–1800.
- Bechara, A. (2001). Neurobiology of decision-making: Risk and reward. *Seminars in Clinical Neuropsychiatry*, 6(3), 205–216.
- Bechara, A. (2003). Risky business: Emotion, decision-making, and addiction. *Journal of Gambling Studies*, 19(1), 23–51.
- Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, 52(2), 336–372.
- Bechara, A., & Damasio, H. (2002). Decision-making and addiction (part I): Impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia*, 40(10), 1675–1689.
- Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10(3), 295–307.
- Bechara, A., Damasio, H., & Damasio, A. R. (2003). Role of the amygdala in decision-making. *Annals of NY Academy of Sciences*, 985, 356–369.
- Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *Journal of Neuroscience*, 19(13), 5473–5481.
- Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience*, 18(1), 428–437.

- Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304), 1293–1295.
- Bechara, A., Tranel, D., & Damasio, H. (2000a). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain*, 123(Pt. 11), 2189–2202.
- De Bellis, M. D., Casey, B. J., Dahl, R. E., Birmaher, B., Williamson, D. E., Thomas, K. M., et al (2000). A pilot study of amygdala volumes in pediatric generalized anxiety disorder. *Biological Psychiatry*, 48(1), 51–57.
- Berridge, K. C. (2007). The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology (Berl)*, 191(3), 391–431.
- Berridge, K. C., & Aldridge, J. W. (2008). *Decision utility, incentive salience, and cue-triggered “wanting”*. The psychology of action. New York: Oxford University Press.
- Birbaumer, N., Grodd, W., Diedrich, O., Klose, U., Erb, M., Lotze, M., et al (1998). fMRI reveals amygdala activation to human faces in social phobias. *Neuroreport*, 9(6), 1223–1226.
- Blair, K. S., Finger, E., Marsh, A. A., Morton, J., Mondillo, K., Buzas, B., et al (2008). The role of 5-HTTLPR in choosing the lesser of two evils, the better of two goods: Examining the impact of 5-HTTLPR genotype and tryptophan depletion in object choice. *Psychopharmacology (Berl)*, 196(1), 29–38.
- Boettiger, C. A., Mitchell, J. M., Tavares, V. C., Robertson, M., Joslyn, G., D'Esposito, M., et al (2007). Immediate reward bias in humans: Fronto-parietal networks and a role for the catechol-o-methyltransferase 158(val/val) genotype. *Journal of Neuroscience*, 27(52), 14383–14391.
- Booij, L., Van der Does, A. J., & Riedel, W. J. (2003). Monoamine depletion in psychiatric and healthy populations: Review. *Molecular Psychiatry*, 8(12), 951–973.
- Borras, L., & Huguelet, P. (2007). Schizophrenia and pathological gambling. *American Journal of Addictions*, 16(4), 269–271.
- Burns, B. D., & Corpus, B. (2004). Randomness and inductions from streaks: “Gambler's fallacy” versus “hot hand”. *Psychonomic Bulletin Review*, 11(1), 179–184.
- Cabrini, S., Baratti, M., Bonfà, F., Cabri, G., Uber, E., & Avanzi, M. (2009). Preliminary evaluation of the DDS-PC inventory: A new tool to assess impulsive-compulsive behaviours associated to dopamine replacement therapy in Parkinson's disease. *Neurological Sciences*.
- Canli, T., Omura, K., Haas, B. W., Fallgatter, A., Constable, R. T., & Lesch, K. P. (2005). Beyond affect: A role for genetic variation of the serotonin transporter in neural activation during a cognitive attention task. *Proceedings of the National Academy of Sciences USA*, 102(34), 12224–12229.
- Cardinal, R. N., Winstanley, C. A., Robbins, T. W., & Everitt, B. J. (2004). Limbic corticostriatal systems and delayed reinforcement. *Annals of NY Academy of Sciences*, 1021, 33–50.
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., et al (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297, 851–854.
- Cavedini, P., Riboldi, G., Keller, R., D'Annunzi, A., & Bellodi, L. (2002). Frontal lobe dysfunction in pathological gambling patients. *Biological Psychiatry*, 51(4), 334–341.
- Charney, D. S. (2003). Neuroanatomical circuits modulating fear and anxiety behaviors. *Acta Psychiatrica Scandinavica*, 108(Suppl. 417), 38–50.
- Cilia, R., Siri, C., Marotta, G., Isaías, I. U., De Gaspari, D., Canesi, M., et al (2008). Functional abnormalities underlying pathological gambling in Parkinson disease. *Archives of Neurology*, 65(12), 1604–1611.
- Clark, L., Bechara, A., Damasio, H., Aitken, M. R., Sahakian, B. J., & Robbins, T. W. (2008). Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain*, 131(Pt. 5), 1311–1322.
- Cools, R., Blackwell, A., Clark, L., Menzies, L., Cox, S., & Robbins, T. W. (2005a). Tryptophan depletion disrupts the motivational guidance of goal-directed behavior as a function of trait impulsivity. *Neuropsychopharmacology*, 30(7), 1362–1373.
- Cools, R., Calder, A. J., Lawrence, A. D., Clark, L., Bullmore, E., & Robbins, T. W. (2005b). Individual differences in threat sensitivity predict serotonergic modulation of amygdala response to fearful faces. *Psychopharmacology (Berl)*, 180(4), 670–679.
- Cooper, J. C., & Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *Neuroimage*, 39(1), 538–547.
- Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 351(1346), 1413–1420.
- Dannon, P. N., Lowengrub, K., Gonopolski, Y., Musin, E., & Kotler, M. (2005). Topiramate versus fluvoxamine in the treatment of pathological gambling: A randomized, blind-rater comparison study. *Clinical Neuropharmacology*, 28(1), 6–10.
- De Luca, V., Tharmalingam, S., Sicard, T., & Kennedy, J. L. (2005). Gene–gene interaction between MAOA and COMT in suicidal behavior. *Neuroscience Letters*, 383, 151–154.
- Desai, R. A., & Potenza, M. N. (2009). A cross-sectional study of problem and pathological gambling in patients with schizophrenia/schizoaffective disorder. *Journal of Clinical Psychiatry*, 70(9), 1250–1257.
- Diatchenko, L., Slade, G. D., Nackley, A. G., Bhalang, K., Sigurdsson, A., Belfer, I., et al (2005). Genetic basis for individual variations in pain perception and the development of a chronic pain condition. *Human Molecular Genetics*, 14(1), 135–143.
- Doornbos, B., Dijck-Brouwer, D. A., Kema, I. P., Tanke, M. A., van Goor, S. A., Muskiet, F. A., et al (2009). The development of peripartum depressive symptoms is associated with gene polymorphisms of MAOA, 5-HTT and COMT. *Progress in Neuropsychopharmacology and Biological Psychiatry*, 33, 1250–1254.
- Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews*, 30(2), 239–271.
- Enoch, M. A., White, K. V., Harris, C. R., Rohrbaugh, J. W., & Goldman, D. (2002). The relationship between two intermediate phenotypes for alcoholism: Low voltage alpha EEG and low P300 ERP amplitude. *Journal of Studies on Alcohol*, 63(5), 509–517.
- Erev, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychology Review*, 105(2), 280–298.
- Erritzøe, D. F., Frokjaer, V. B., Christoffersen, M. V., Baare, W., Ramsoy, T., Svarer, C., et al (2008). Decreased serotonin-2a binding in MDMA and hallucinogen users: An [¹⁸F] altanserin PET study. *Neuroimage*, 41, 46.
- Evans, A. H., Strafella, A. P., Weintraub, D., & Stacy, M. (2009). Impulsive and compulsive behaviors in Parkinson's disease. *Movement Disorders*.
- Evers, E. A., Cools, R., Clark, L., van der Veen, F. M., Jolles, J., Sahakian, B. J., et al (2005). Serotonergic modulation of prefrontal cortex during negative feedback in probabilistic reversal learning. *Neuropsychopharmacology*, 30(6), 1138–1147.
- Fellows, L. K. (2007). Advances in understanding ventromedial prefrontal function: The accountant joins the executive. *Neurology*, 68(13), 991–995.
- Ferrara, J. M., & Stacy, M. (2008). Impulse–Control disorders in Parkinson's disease. *CNS Spectrums*, 13(8), 690–698.
- Figgitt, D. P., & McClellan, K. J. (2000). Fluvoxamine. An updated review of its use in the management of adults with anxiety disorders. *Drugs*, 60(4), 925–954.
- Finger, E. C., Marsh, A. A., Buzas, B., Kamel, N., Rhodes, R., Vytlachil, M., et al (2007). The impact of tryptophan depletion and 5-HTTLPR genotype on passive avoidance and response reversal instrumental learning tasks. *Neuropsychopharmacology*, 32(1), 206–215.
- Fischer, H., Andersson, J. L., Furmark, T., & Fredrikson, M. (2000). Fear conditioning and brain activity: A positron emission tomography study in humans. *Behavioral Neuroscience*, 114(4), 671–680.
- Foti, D., & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards: Evidence from event-related potentials. *Biological Psychology*, 87(1), 1–8.
- Frith, C., & Dolan, R. (1996). The role of the prefrontal cortex in higher cognitive functions. *Brain Research. Cognitive Brain Research*, 5(1–2), 175–181.
- Gallhofer, B., Lis, S., Meyer-Lindenberg, A., & Krieger, S. (1999). Cognitive dysfunction in schizophrenia: A new set of tools for the assessment of cognition and drug effects. *Acta Psychiatrica Scandinavica Supplementum*, 395, 118–128.
- Goldberg, T. E., Egan, M. F., Gscheidle, T., Coppola, R., Weickert, T., Kolachana, B. S., et al (2003). Executive subprocesses in working memory: Relationship to catechol-o-methyltransferase val158met genotype and schizophrenia. *Archives of General Psychiatry*, 60(9), 889–896.
- Gorman, J. M., & Kent, J. M. (1999). SSRIs and SNRIs: Broad spectrum of efficacy beyond major depression. *Journal of Clinical Psychiatry*, 60(Suppl. 4), 33–38 [discussion 39].

- Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry*, 160(4), 636–645.
- Hall, G. W., Carrero, N. J., Takushi, R. Y., Montoya, I. D., Preston, K. L., & Gorelick, D. A. (2000). Pathological gambling among cocaine-dependent outpatients. *American Journal of Psychiatry*, 157(7), 1127–1133.
- Hariri, A. R., & Holmes, A. (2006). Genetics of emotional regulation: The role of the serotonin transporter in neural function. *Trends in Cognitive Sciences*, 10(4), 182–191.
- Hariri, A. R., Drabant, E. M., & Weinberger, D. R. (2006). Imaging genetics: Perspectives from studies of genetically driven variation in serotonin function and corticolimbic affective processing. *Biological Psychiatry*, 59(10), 888–897.
- Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., et al (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science*, 297(5580), 400–403.
- Harmer, C. J., Rogers, R. D., Tunbridge, E., Cowen, P. J., & Goodwin, G. M. (2003). Tryptophan depletion decreases the recognition of fear in female volunteers. *Psychopharmacology (Berl)*, 167(4), 411–417.
- Harris, G. C., & Aston-Jones, G. (2006). Arousal and reward: A dichotomy in orexin function. *Trends in Neurosciences*, 29, 571–577.
- Heerey, E. A., Bell-Warren, K. R., & Gold, J. M. (2008). Decision-Making impairments in the context of intact reward sensitivity in schizophrenia. *Biological Psychiatry*, 64(1), 62–69.
- Hollander, E., DeCaria, C. M., Finkell, J. N., Begaz, T., Wong, C. M., & Cartwright, C. (2000). A randomized double-blind fluvoxamine/placebo crossover trial in pathologic gambling. *Biological Psychiatry*, 47(9), 813–817.
- Homberg, J. R., van den Bos, R., den Heijer, E., Suer, R., & Cuppen, E. (2008). Serotonin transporter dosage modulates long-term decision-making in rat and human. *Neuropharmacology*, 55(1), 80–84.
- Hurlemann, R., Schlaepfer, T. E., Matusch, A., Reich, H., Shah, N. J., Zilles, K., et al (2009). Reduced 5-HT(2A) receptor signaling following selective bilateral amygdala damage. *Social Cognitive and Affective Neuroscience*, 4(1), 79–84.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, XLVII, 263–291.
- Kahneman, D. (1994). New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics*, 150, 18–36.
- Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112, 375–405.
- Kalbitzer, J., Frokjaer, V. G., Erritzoe, D., Svarer, C., Cumming, P., Nielsen, F. A., et al (2009). The personality trait openness is related to cerebral 5-HT levels. *Neuroimage*, 45(2), 280–285.
- Kähkänen, S., Mäkinen, V., Jääskeläinen, I. P., Pennanen, S., Liesivuori, J., & Ahveninen, J. (2005). Serotonergic modulation of mismatch negativity. *Psychiatry Research*, 138(1), 61–74.
- Kirk, U. (2008). The neural basis of object–context relationships on aesthetic judgment. *PloS ONE*, 3(11), e3754.
- Kirk, U., Skov, M., Hulme, O., Christensen, M. S., & Zeki, S. (2009). Modulation of aesthetic value by semantic context: An fmri study. *Neuroimage*, 44(3), 1125–1132.
- Knutson, B., & Bossaerts, P. (2007). Neural antecedents of financial decisions. *Journal of Neuroscience*, 27(31), 8174–8177.
- Knutson, B., Rick, S., Wimmer, G. E., Prelec, D., & Loewenstein, G. (2007). Neural predictors of purchases. *Neuron*, 53(1), 147–156.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al (2007). Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*, 446(7138), 908–911.
- Kreek, M. J., Nielsen, D. A., Butelman, E. R., & LaForge, K. S. (2005). Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nature Neuroscience*, 8(11), 1450–1457.
- Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk taking. *Neuron*, 47(5), 763–770.
- Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., et al (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, 35(11), 1437–1444.
- Launzenberger, R. R., Mitterhauser, M., Spindelberger, C., Wadsak, W., Klein, N., Mien, L. K., et al (2007). Reduced serotonin-1a receptor binding in social anxiety disorder. *Biological Psychiatry*, 61(9), 1081–1089.
- LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behavioral Brain Research*, 58(1–2), 69–79.
- LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cell Molecular Neurobiology*, 23(4–5), 727–738.
- Levita, L., Hare, T. A., Voss, H. U., Glover, G., Ballon, D. J., & Casey, B. J. (2009). The bivalent side of the nucleus accumbens. *Neuroimage*, 44(3), 1178–1187.
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychology Bulletin*, 127(2), 267–286.
- MacDonald, S., Mann, R., Chipman, M., Pakula, B., Erickson, P., Hathaway, A., et al (2008). Driving behavior under the influence of cannabis or cocaine. *Traffic Injury Prevention*, 9(3), 190–194.
- Marsh, A. A., Finger, E. C., Buzas, B., Soliman, N., Richell, R. A., Vytlachil, M., et al (2006). Impaired recognition of fear facial expressions in 5-HTLPRs-polymorphism carriers following tryptophan depletion. *Psychopharmacology (Berl)*, 189(3), 387–394.
- McClure, S. M., Daw, N. D., & Montague, P. R. (2003). A computational substrate for incentive salience. *Trends in Neurosciences*, 26(8), 423–428.
- McClure, S. M., Li, J., Tomlin, D., Cybert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44(2), 379–387.
- Merens, W., Willem Van Der Does, A. J., & Spinhoven, P. (2007). The effects of serotonin manipulations on emotional information processing and mood. *Journal of Affective Disorders*, 103(1–3), 43–62.
- Merims, D., & Giladi, N. (2008). Dopamine dysregulation syndrome, addiction and behavioral changes in Parkinson's disease. *Parkinsonism and Related Disorders*, 14(4), 273–280.
- Meyer, J. H., Houle, S., Sagrati, S., Carella, A., Hussey, D. F., Ginovart, N., et al (2004). Brain serotonin transporter binding potential measured with carbon 11-labeled DASB positron emission tomography: Effects of major depressive episodes and severity of dysfunctional attitudes. *Archives of General Psychiatry*, 61(12), 1271–1279.
- Meyer, J. H., Kapur, S., Eisfeld, B., Brown, G. M., Houle, S., DaSilva, J., et al (2001). The effect of paroxetine on 5-HT(2A) receptors in depression: An [(18)F]setoperone PET imaging study. *American Journal of Psychiatry*, 158(1), 78–85.
- Meyer-Lindenberg, A., Buckholtz, J. W., Kolachana, B., Hariri, A. R., Pezawas, L., Blasi, G., et al (2006). Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proceedings of the National Academy of Sciences USA*, 103, 6269–6274.
- Meyer-Lindenberg, A., & Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience*, 7(10), 818–827.
- Meyer-Lindenberg, A., Kohn, P. D., Kolachana, B., Kippenhan, S., McLnerney-Leo, A., Nussbaum, R., et al (2005). Midbrain dopamine and prefrontal function in humans: Interaction and modulation by COMT genotype. *Nature Neuroscience*, 8(5), 594–596.
- Meyer-Lindenberg, A., Miletich, R. S., Kohn, P. D., Esposito, G., Carson, R. E., Quarantelli, M., et al (2002). Reduced prefrontal activity predicts exaggerated striatal dopaminergic function in schizophrenia. *Nature Neuroscience*, 5(3), 267–271.
- Meyer-Lindenberg, A., Nichols, T., Callicott, J. H., Ding, J., Kolachana, B., Buckholtz, J., et al (2006). Impact of complex genetic variation in COMT on human brain function. *Molecular Psychiatry*, 11(9), 867–877, 797.
- Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., & Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addiction*, 96(12), 1825–1837.
- Montoya, A. G., Sorrentino, R., Lukas, S. E., & Price, B. H. (2002). Long-Term neuropsychiatric consequences of "ecstasy" (MDMA): A review. *Harvard Review of Psychiatry*, 10(4), 212–220.
- Morton, J. (2005). Ecstasy: Pharmacology and neurotoxicity. *Current Opinion in Pharmacology*, 5(1), 79–86.
- Munafò, M. R., Hayward, G., & Harmer, C. (2006). Selective processing of social threat cues following acute tryptophan depletion. *Journal of Psychopharmacology*, 20(1), 33–39.

- Murray, E. A. (2007). The amygdala, reward and emotion. *Trends in Cognitive Sciences*, 11(11), 489–497.
- Nash, J. R., Sargent, P. A., Rabiner, E. A., Hood, S. D., Argyropoulos, S. V., Potokar, J. P., et al (2008). Serotonin 5-HT1A receptor binding in people with panic disorder: Positron emission tomography study. *British Journal of Psychiatry*, 193, 229–234.
- Neumeister, A., Konstantinidis, A., Stastny, J., Schwarz, M. J., Vitouch, O., Willeit, M., et al (2002). Association between serotonin transporter gene promoter polymorphism (5HTTLPR) and behavioral responses to tryptophan depletion in healthy women with and without family history of depression. *Archives of General Psychiatry*, 59(7), 613–620.
- Nugent, A. C., Neumeister, A., Goldman, D., Herscovitch, P., Charney, D. S., & Drevets, W. C. (2008). Serotonin transporter genotype and depressive phenotype determination by discriminant analysis of glucose metabolism under acute tryptophan depletion. *Neuroimage*, 43(4), 764–774.
- O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, 33, 815–826.
- Ondo, W. G., & Lai, D. (2008). Predictors of impulsivity and reward seeking behavior with dopamine agonists. *Parkinsonism Related Disorders*, 14(1), 28–32.
- Parrott, A. C. (2001). Human psychopharmacology of ecstasy (MDMA): A review of 15 years of empirical research. *Human Psychopharmacology*, 16(8), 557–577.
- Pattij, T., & Vandervschuren, L. J. (2008). The neuropharmacology of impulsive behaviour. *Trends in Pharmacological Sciences*, 29(4), 192–199.
- Paulus, M. P., Feinstein, J. S., Simmons, A., & Stein, M. B. (2004). Anterior cingulate activation in high trait anxious subjects is related to altered error processing during decision making. *Biological Psychiatry*, 55(12), 1179–1187.
- Peciña, S., & Berridge, K. (2005). Hedonic hot spot in nucleus accumbens shell: Where do mu-opioids cause increased hedonic impact of sweetness? *Journal of Neuroscience*, 25, 11777–11786.
- Peciña, S., Smith, K. S., & Berridge, K. C. (2006). Hedonic hot spots in the brain. *Neuroscientist*, 12, 500–511.
- Peterson, R. L. (2007). *Inside the investor's brain: The power of mind over money*. Wiley Inc.
- Petrovic, P., Plegier, B., Seymour, B., Klöppel, S., De Martino, B., Critchley, H., et al (2008). Blocking central opiate function modulates hedonic impact and anterior cingulate response to rewards and losses. *Journal of Neuroscience*, 28, 10509–10516.
- Pfau, J. G. (2009). Pathways of sexual desire. *Journal of Sexual Medicine*.
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of National Academy of Science USA*, 105(3), 1050–1054.
- Praschak-Rieder, N., Wilson, A. A., Hussey, D., Carella, A., Wei, C., Ginovart, N., et al (2005). Effects of tryptophan depletion on the serotonin transporter in healthy humans. *Biological Psychiatry*, 58(10), 825–830.
- Price, J. L. (1999). Prefrontal cortical networks related to visceral function and mood. *Annals of NY Academy of Sciences*, 877, 383–396.
- Quednow, B. B., Kühn, K. U., Hoppe, C., Westheide, J., Maier, W., Daum, I., et al (2007). Elevated impulsivity and impaired decision-making cognition in heavy users of MDMA ("Ecstasy"). *Psychopharmacology (Berl)*, 189, 517–530.
- Ramsøy, T. Z., Madsen, K. H., Wegener, J. S., Gelskov, S. A. V., Erritzøe, D., Knudsen, G. M., et al (2009a). From ecstasy to agony: Chronic effects of MDMA use on emotional processing [abstract]. *Human Brain Mapping Annual Conference*.
- Ramsøy, T. Z., Madsen, K. S., Madsen, K. H., Wegener, J. S., Knudsen, G. M., Jernigan, T. L., et al (2009b). Effects of recreational use of ecstasy and hallucinogens on the neural circuitry of emotion [abstract]. *Society for Neuroscience Annual Conference*.
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, 9(7), 545–556.
- Rauch, S. L., Shin, L. M., & Wright, C. I. (2003). Neuroimaging studies of amygdala function in anxiety disorders. *Annals NY Academy of Science*, 985, 389–410.
- Reimold, M., Batra, A., Knobel, A., Smolka, M. N., Zimmer, A., Mann, K., et al (2008). Anxiety is associated with reduced central serotonin transporter availability in unmedicated patients with unipolar major depression: A [11C]DASB PET study. *Molecular Psychiatry*, 13(6), 606–613, 557.
- Rey, A., Goldstein, R. M., & Perruchet, P. (2009). Does unconscious thought improve complex decision making? *Psychological Research*, 73, 372–379.
- Riedel, W. J., Klaassen, T., & Schmitt, J. A. (2002). Tryptophan, mood, and cognitive function. *Brain Behaviour and Immunity*, 16(5), 581–589.
- Rogers, R. D., Tunbridge, E. M., Bhagwagar, Z., Drevets, W. C., Sahakian, B. J., & Carter, C. S. (2003). Tryptophan depletion alters the decision-making of healthy volunteers through altered processing of reward cues. *Neuropsychopharmacology*, 28(1), 153–162.
- Roiser, J. P., de Martino, B., Tan, G. C., Kumaran, D., Seymour, B., Wood, N. W., et al (2009). A genetically mediated bias in decision making driven by failure of amygdala control. *Journal of Neuroscience*, 29(18), 5985–5991.
- Roiser, J. P., Müller, U., Clark, L., & Sahakian, B. J. (2007). The effects of acute tryptophan depletion and serotonin transporter polymorphism on emotional processing in memory and attention. *International Journal of Neuropsychopharmacology*, 10(4), 449–461.
- Roussos, P., Giakoumaki, S. G., Pavlakis, S., & Bitsios, P. (2008). Planning, decision-making and the COMT rs4818 polymorphism in healthy males. *Neuropsychologia*, 46(2), 757–763.
- Royer, J. P., Zald, D., Versace, R., Costes, N., Lavenne, F., Koenig, O., et al (2000). Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: A positron emission tomography study. *Journal of Neuroscience*, 20(20), 7752–7759.
- Ruhé, H. G., Mason, N. S., & Schene, A. H. (2007). Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: A meta-analysis of monoamine depletion studies. *Molecular Psychiatry*, 12(4), 331–359.
- Sanfey, A. G., & Chang, L. J. (2008). Multiple systems in decision making. *Annals of NY Academy of Sciences*, 1128, 53–62.
- Sanfey, A. G., Loewenstein, G., McClure, S. M., & Cohen, J. D. (2006). Neuroeconomics: Cross-currents in research on decision-making. *Trends in Cognitive Science*, 10(3), 108–116.
- Savitz, J., Lucki, I., & Drevets, W. C. (2009). 5-HT(1A) receptor function in major depressive disorder. *Progress in Neurobiology*, 88(1), 17–31.
- Schott, B. H., Minuzzi, L., Krebs, R. M., Elmenhorst, D., Lang, M., Winz, O. H., et al (2008). Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *Journal of Neuroscience*, 28(52), 14311–14319.
- Sevy, S., Hassoun, Y., Bechara, A., Yechiam, E., Napolitano, B., Burdick, K., et al (2006). Emotion-based decision-making in healthy subjects: Short-term effects of reducing dopamine levels. *Psychopharmacology (Berl)*, 188(2), 228–235.
- Shiv, B., Loewenstein, G., Bechara, A., Damasio, H., & Damasio, A. (2005). Investment behavior and the negative side of emotion. *Psychological Science*, 16(6), 435–439.
- Smoski, M. J., Lynch, T. R., Rosenthal, M. Z., Cheavens, J. S., Chapman, A. L., & Krishnan, R. R. (2008). Decision-making and risk aversion among depressive adults. *Journal of Behavior Therapy Experimental Psychiatry*, 39(4), 567–576.
- Sokolski, K. N., Conney, J. C., Brown, B. J., & DeMet, E. M. (2004). Once-daily high-dose pindolol for ssri-refractory depression. *Psychiatry Research*, 125(2), 81–86.
- Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11, 543–545.
- Spillmann, M. K., Van der Does, A. J., Rankin, M. A., Vuolo, R. D., Alpert, J. E., Nierenberg, A. A., et al (2001). Tryptophan depletion in ssri-recovered depressed outpatients. *Psychopharmacology (Berl)*, 155(2), 123–127.
- Stalnaker, T. A., Roesch, M. R., Franz, T. M., Calu, D. J., Singh, T., & Schoenbaum, G. (2007). Cocaine-induced decision-making deficits are mediated by miscoding in basolateral amygdala. *Nature Neuroscience*, 10(8), 949–951.
- Stamey, W., & Jankovic, J. (2008). Impulse control disorders and pathological gambling in patients with Parkinson disease. *Neurologist*, 14(2), 89–99.
- Steeves, T. D., Miyasaki, J., Zurowski, M., Lang, A. E., Pellecchia, G., Van Eimeren, T., et al (2009). Increased striatal dopamine release in parkinsonian patients with pathological gambling: A [11C] raclopride PET study. *Brain*, 132(Pt 5), 1376–1385.
- Stein, M. B., Goldin, P. R., Sareen, J., Zorrilla, L. T., & Brown, G. G. (2002). Increased amygdala activation to angry and contemptuous faces in generalized social phobia. *Archives of General Psychiatry*, 59(11), 1027–1034.
- Strauss, J., Barr, C. L., George, C. J., Ryan, C. M., King, N., Shaikh, S., et al (2004). BDNF and COMT polymorphisms: Relation to memory phenotypes in young adults with childhood-onset mood disorder. *Neuromolecular Medicine*, 5(3), 181–192.
- Suhler, C. L., & Churchland, P. S. (2009). Control: conscious and otherwise. *Trends in Cognitive Science*, 13, 341–347.

- Surguladze, S. A., Elkin, A., Ecker, C., Kalidindi, S., Corsico, A., Giampietro, V., et al (2008). Genetic variation in the serotonin transporter modulates neural system-wide response to fearful faces. *Genes Brain Behaviour*, 7(5), 543–551.
- Suri, R. E. (2002). TD models of reward predictive responses in dopamine neurons. *Neural Networks*, 15, 523–533.
- Takahashi, T. (2007). Hyperbolic discounting may be reduced to electrical coupling in dopaminergic neural circuits. *Medical Hypotheses*, 69(1), 195–198.
- Tan, H. Y., Chen, Q., Goldberg, T. E., Mattay, V. S., Meyer-Lindenberg, A., Weinberger, D. R., et al (2007). Catechol-o-methyltransferase val158met modulation of prefrontal-parietal-striatal brain systems during arithmetic and temporal transformations in working memory. *Journal of Neuroscience*, 27(49), 13393–13401.
- Tang, X., Guo, S., Sun, H., Song, X., Jiang, Z., Sheng, L., et al (2009). Gene–gene interactions of CYP2A6 and MAOA polymorphisms on smoking behavior in Chinese male population. *Pharmacogenetics and Genomics*, 19, 345–352.
- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, 315(5811), 515–518.
- Torta, D. M., & Castelli, L. (2008). Reward pathways in Parkinson's disease: Clinical and theoretical implications. *Psychiatry and Clinical Neuroscience*, 62(2), 203–213.
- van der Sluis, S., Willemsen, G., de Geus, E. J., Boomsma, D. I., & Posthumus, D. (2008). Gene–environment interaction in adults' IQ scores: Measures of past and present environment. *Behaviour Genetics*, 38, 348–360.
- van 't Wout, M., Kahn, R. S., Sanfey, A. G., & Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Experimental Brain Research*, 169(4), 564–568.
- Volavka, J., Bildner, R., & Nolan, K. (2004). Catecholamines and aggression: The role of COMT and MAO polymorphisms. *Annals of NY Academy of Sciences*, 1036, 393–398.
- Volkow, N. D., Fowler, J. S., Wang, G. J., Baler, R., & Telang, F. (2009). Imaging Dopamine's role in drug abuse and addiction. *Neuropharmacology*, 56(Suppl 1), 3–8.
- Volkow, N. D., Fowler, J. S., Wang, G. J., Swanson, J. M., & Telang, F. (2007). Dopamine in drug abuse and addiction: Results of imaging studies and treatment implications. *Archives of Neurology*, 64(11), 1575–1579.
- Verholt, V. (2008). The neuronal substrate of risky choice: An insight into the contributions of neuroimaging to the understanding of theories on decision making under risk. *Annals of NY Academy of Sciences*, 1128, 41–52.
- Vuilleumier, P., Armony, J. L., Clarke, K., Husain, M., Driver, J., & Dolan, R. J. (2002). Neural response to emotional faces with and without awareness: Event-related fmri in a parietal patient with visual extinction and spatial neglect. *Neuropsychologia*, 40(12), 2156–2166.
- Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 11, 1271–1278.
- Watson, K. K., Ghodasra, J. H., & Platt, M. L. (2009). Serotonin transporter genotype modulates social reward and punishment in rhesus macaques. *Plos ONE*, 4(1), e4156.
- Weber, E. U., & Johnson, E. J. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, 60, 53–85.
- Williams, L. M., Gatt, J. M., Schofield, P. R., Olivieri, G., Peduto, A., & Gordon, E. (2009). 'Negativity bias' in risk for depression and anxiety: Brain-body fear circuitry correlates, 5-HTT-LPR and early life stress. *Neuroimage*, 47(3), 804–814.
- de Win, M. M., Reneman, L., Reitsma, J. B., den Heeten, G. J., Booij, J., & van den Brink, W. (2004). Mood disorders and serotonin transporter density in ecstasy users – the influence of long-term abstention, dose, and gender. *Psychopharmacology (Berl)*, 173(3–4), 376–382.
- Winterer, G., Hariri, A. R., Goldman, D., & Weinberger, D. R. (2005). Neuroimaging and human genetics. *International Review of Neurobiology*, 67, 325–383.
- Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience*, 26(37), 9530–9537.
- Yatham, L. N., Liddle, P. F., Shah, I. S., Lam, R. W., Adam, M. J., Zis, A. P., et al (2001). Effects of rapid tryptophan depletion on brain 5-HT(2) receptors: A PET study. *British Journal of Psychiatry*, 178, 448–453.
- Yip, S. W., Sacco, K. A., George, T. P., & Potenza, M. N. (2009). Risk/reward decision-making in schizophrenia: A preliminary examination of the influence of tobacco smoking and relationship to Wisconsin card sorting task performance. *Schizophrenia Research*, 110(1–3), 156–164.
- Young, S. N., & Leyton, M. (2002). The role of serotonin in human mood and social interaction. Insight from altered tryptophan levels. *Pharmacology Biochemistry and Behaviour*, 71(4), 857–865.
- Zak, P. J. (2004). Neuroeconomics. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 359(1451), 1737–1748.
- Zak, P. J., & Fakhari, A. (2006). Neuroactive hormones and interpersonal trust: International evidence. *Economics and Human Biology*, 4(3), 412–429.
- Zak, P. J., Kurzban, R., & Matzner, W. T. (2004). The neurobiology of trust. *Annals of NY Academy of Sciences*, 1032, 224–227.
- Zubieta, J. K., Heitzeg, M. M., Smith, Y. R., Bueller, J. A., Xu, K., Xu, Y., et al (2003). COMT val158met genotype affects mu-opioid neurotransmitter responses to a pain stressor. *Science*, 299(5610), 1240–1243.
- Zuckerman, M., Ballenger, J. C., & Post, R. M. (1984). The neurobiology of some dimensions of personality. *International Review of Neurobiology*, 25, 391–436.

SOCIAL BEHAVIOR



Social factors are well known to affect our choices and behaviors as consumers. We are affected by peers and culture. But how does this exactly occur? What are the exact mechanisms with which “social thinking” affects what we do?

One part of consumer neuroscience leans towards social psychology and social neuroscience, and recent studies provide a better understanding of how social contexts can affect our choices.

Social behavior, as so many other mental operations, cover a range of different processes and functions. The mere fact that we are so social beings reflects a fundamental fact about humans, and the way in which our brains are wired to be social from the get-go cannot sufficiently emphasize the fact that social brains is one of the new frontiers in neuromarketing.



How social cognition can inform social decision making

Victoria K. Lee^{1*} and Lasana T. Harris^{1,2}

¹ Department of Psychology and Neuroscience, Duke University, Durham, NC, USA

² Center for Cognitive Neuroscience, Duke University, Durham, NC, USA

Edited by:

Masaki Isoda, Kansai Medical University, Japan

Reviewed by:

V. S. Chandrasekhar Pammi,
University of Allahabad, India
Shinsuke Suzuki, California Institute of Technology, USA

***Correspondence:**

Victoria K. Lee, Department of Psychology and Neuroscience, Duke University, Box 90086, 417 Chapel Drive, Durham, NC 27708-0086, USA
e-mail: vkl3@duke.edu

Social decision-making is often complex, requiring the decision-maker to make inferences of others' mental states in addition to engaging traditional decision-making processes like valuation and reward processing. A growing body of research in neuroeconomics has examined decision-making involving social and non-social stimuli to explore activity in brain regions such as the striatum and prefrontal cortex, largely ignoring the power of the social context. Perhaps more complex processes may influence decision-making in social vs. non-social contexts. Years of social psychology and social neuroscience research have documented a multitude of processes (e.g., mental state inferences, impression formation, spontaneous trait inferences) that occur upon viewing another person. These processes rely on a network of brain regions including medial prefrontal cortex (MPFC), superior temporal sulcus (STS), temporal parietal junction, and precuneus among others. Undoubtedly, these social cognition processes affect social decision-making since mental state inferences occur spontaneously and automatically. Few studies have looked at how these social inference processes affect decision-making in a social context despite the capability of these inferences to serve as predictions that can guide future decision-making. Here we review and integrate the person perception and decision-making literatures to understand how social cognition can inform the study of social decision-making in a way that is consistent with both literatures. We identify gaps in both literatures—while behavioral economics largely ignores social processes that spontaneously occur upon viewing another person, social psychology has largely failed to talk about the implications of social cognition processes in an economic decision-making context—and examine the benefits of integrating social psychological theory with behavioral economic theory.

Keywords: social cognition, person perception, social decision-making, economic games, computers

What makes social decision-making unique and different from non-social decision-making? Humans are highly social animals—as such, researchers often take for granted the ease with which humans make social decisions. This begs the question whether social decision-making is a simplified type of decision-making. Yet social decision-making should be a complex process—social decision-makers must engage traditional decision-making processes (e.g., learning, valuation, and feedback processing), as well as infer the mental states of another person. These two tasks have been separately studied in the fields of behavioral economics and social psychology, with behavioral economists studying decision-making in interactive economic games and social psychologists studying spontaneous inferences about other people. Each of these fields has separately made major contributions to the understanding of social behavior. However, a more cohesive theory of social decision-making results when researchers combine these literatures.

When talking about social decision-making, many different types of decisions may come to mind—decisions about other people (Is Linda a feminist bank teller?), decisions that are influenced by other people (e.g., social conformity and expert advice), as well as decisions that are interactive (e.g., two people want to go to dinner but have to decide on a restaurant). In this review,

we focus on strategic interaction decisions often employed in behavioral economics games (e.g., trust game, ultimatum game, prisoner's dilemma game, etc.) that require thinking about the mental states of another person. Research shows that such decisions may differ depending on whether the interaction partner is another person or a computer agent. Here, we suggest that such differences in decision-making arise due to differences when processing human and computer agents. Specifically, viewing another person engages the social cognition brain network, allowing for mental state inferences that function as predictions during the decision phase, as well as spontaneous trait inferences that occur when viewing the other person's behavior in the feedback phase.

To understand how decision-making in a social context is different than non-social decision-making, it is first important to understand what exactly makes humans unique as social agents. Social psychological theory suggests humans differ from objects in important ways (Fiske and Taylor, 2013). First, humans are intentional agents that influence and try to control the environment for their own purposes. Computers on the other hand are non-intentional agents. The decisions made by a computer result from fixed, preprogrammed algorithms, and are usually not as flexible as human decision-making. Second, people form impressions of others at the same time others are forming impressions

of them. Therefore, in a social situation people are trying to form impressions of another person at the same time they are trying to manage the impression being formed of them. In meaningful social interaction (most social interactions) the first person usually cares about the reputation the second person is forming of them, wanting them to form a largely positively valenced impression. Each interaction partner is aware that they are the target of someone's attention and may monitor or change their behavior as a result. Third, it is harder to verify the accuracy of one's cognitions about a person than they are about an object. Because things like traits, which are essential to thinking about people, are invisible features of a person and are often inferred, it is harder to verify that a person is trustworthy than it is to verify that a computer, for example, is trustworthy. This may be because the person can manipulate trait information such as trustworthiness—an immoral person can act in moral ways when desired—but a computer has no such desire. Last, and perhaps most importantly, humans possess mental states—thoughts and feelings that presumably cause behavior—that are only known to them. People automatically try to infer the mental states of others because such inferences facilitate social interactions. Computers, however, do not have mental states because they do not have minds. This important distinction—the possession of mental states—allows for the differences mentioned above in intentionality and impression management. These key differences allow us to examine what these social cognitive processes (impression management and intentionality) contribute to the uniqueness of social decision-making, though this discussion seems to often elude studies of social decision-making.

There are also important similarities between humans and computers that make computers the ideal comparison in social decision-making studies. With analogies comparing the human brain to a computer, it almost seems natural that many studies have turned to computers as the non-social comparison. Computers, like humans, are *agents* that can take actions toward a participant. Presumably a computer can "decide" to share money in a trust game as can a human partner. Additionally both humans and computers are information processing systems. Participants' decisions are presumably "registered" by both human and computer agents. Advanced computer programs can take participants' choices into account in order to "learn" to predict another person's behavior using programmed algorithms. For example, website ads learn to predict what a person may purchase based on search history. In some economic games, a computer's responses may be dependent on the participant's past decisions. These similarities allow researchers to compare decisions across agents and examine what social agents add to the decision-making process.

SOCIAL DECISION-MAKING BRAIN REGIONS

One way to understand the unique nature of social decision-making is to take a neuroscientific approach. By understanding what goes on in the brain, we can begin to dissociate social and non-social decisions. This strategy is particularly informative and useful because similar behavior is sometimes observed for social and non-social stimuli, but the neural mechanisms underlying those decisions are found to be different (e.g., Harris et al., 2005; Harris and Fiske, 2008). Below, we briefly summarize two brain

networks we believe will be involved in social decision-making—the traditional decision-making brain network, and the social cognition/person perception brain network¹. As a caveat, the reader must remember when discussing the unique qualities of social decision-making, we are still examining decision-making. As such, traditional decision-making processes and brain structures underlying these processes are involved in social decision-making studies. Past studies demonstrate that the social context modulates these decision-making structures (see Engelmann and Hein, 2013 for review). However, exactly *how* the social context does this is not entirely understood. By looking in the social cognition/person perception brain network, researchers are beginning to explore how these functions are integrated at a neural level (e.g., Hampton et al., 2008; Yoshida et al., 2010; Suzuki et al., 2012). Next, we list brain regions implicated in decision-making and social cognition.

Past research shows decision-making brain regions are also involved in social decision-making. The medial prefrontal cortex (MPFC)—responsible for creating value signals for food, non-food consumables, and monetary gambles (Chib et al., 2009)—is also active when creating value signals in a social context (Lin et al., 2012). These value signals can be thought of as a quantifiable signal for making predictions—those assigned a higher value predict a better outcome, and those assigned a lower value predict a worse outcome. Recently, it has been suggested that the MPFC works as an action-outcome predictor concerned with learning and predicting the likelihood of outcomes associated with actions (Alexander and Brown, 2011). Similarly, investigations of social reward processing suggest that the striatum responds to both social and monetary rewards (Izuma et al., 2008, 2010). The connections between cortical and subcortical regions with the striatum create a network of brain regions engaged during decision-making. The neurotransmitter dopamine provides a vehicle by which these brain regions communicate. Prediction error signals—the firing of dopamine neurons when observed outcomes differ from expectations (or predictions)—also occur for social stimuli in economic games (Lee, 2008; Rilling and Sanfey, 2011) as well as when social targets violate expectations (Harris and Fiske, 2010). Collectively these regions, along with other regions such as the amygdala, posterior cingulate cortex (PCC), insula, and other areas of prefrontal cortex including orbital prefrontal cortex and a more rostral region of MPFC make up a decision-making network often engaged during economic decision-making (Knutson and Cooper, 2005; Delgado et al., 2007).

While social decision-making studies have investigated how the striatum and prefrontal cortex are modulated by the social context, another prevalent question is whether a network of brain regions established in the social neuroscience literature on social cognition and person perception is also active during social decision-making and how these brain regions interact. An important part of social cognition consists of inferring mental

¹However before we begin, it should be noted that it is easy to make these distinctions for discussion purposes here, but each of these processes rely on other brain regions as well and the decision-making process is the result of interactions between these brain regions.

states, like the intentions of a social target (Frith and Frith, 2001). During tasks that involve dispositional attributions—an inference of an enduring mental state—areas such as MPFC and superior temporal sulcus (STS) are reliably activated (Harris et al., 2005). Other areas involved in person perception include temporal-parietal junction (TPJ), pregenual anterior cingulate cortex (pACC), amygdala, insula, fusiform gyrus of temporal cortex (FFA), precuneus, posterior cingulate, temporal pole, and inferior parietal cortex (IPL; Gallese et al., 2004; Haxby et al., 2004; Amadio and Frith, 2006). Together these regions represent a social cognition network that can be used to navigate the social world. This network is believed to be activated in a variety of social cognition tasks, including thinking about others' intentions and goals (i.e., theory of mental state tasks), identifying social others (i.e., faces and bodily movement), moral judgments, social scripts, and making trait inferences (see Van Overwalle, 2009, for a review). However, until recently the mention of these regions in social decision-making studies has been scarce, often being relegated to a supplemental analysis or table. Presumably these social cognitive processes are relevant for decision-making when interacting with human agents because they occur automatically and with minimal exposure to the social target (Ambady and Rosenthal, 1992; Willis and Todorov, 2006). Therefore, these automatic social processes are most likely engaged in a social decision-making context and perhaps provide the vehicle through which the social context modulates decision-making brain regions like the striatum and PFC.

DIFFERENCES IN SOCIAL AND NONSOCIAL DECISION-MAKING PROCESSES

Decision-making in its most basic form can be broken down into three key processes², (1) making predictions that guide decision-making, (2) examining the outcome of the decision, and (3) using the outcome to update predictions, a process often described as learning. Next, we discuss differences between humans and computers for each of these aspects of decision-making to understand how social decision-making is unique (see Figure 1 for a summary of these findings).

Social predictions

Predictions have received much attention when studying social decision-making. Behavioral economics games such as the trust game, ultimatum game, or the prisoner's dilemma game are often used to study social preferences for trustworthiness, fairness, or cooperation, respectively. However, each of these games requires *predicting* what another agent (person or computer) will do. The combination of the participant's and the partner's decisions determines the outcome. Therefore, in order to maximize payout, the participant has to predict what the partner will do and decide accordingly. What information do participants rely on when making these predictions? Social psychological theory suggests these predictions rely on trait inferences that occur

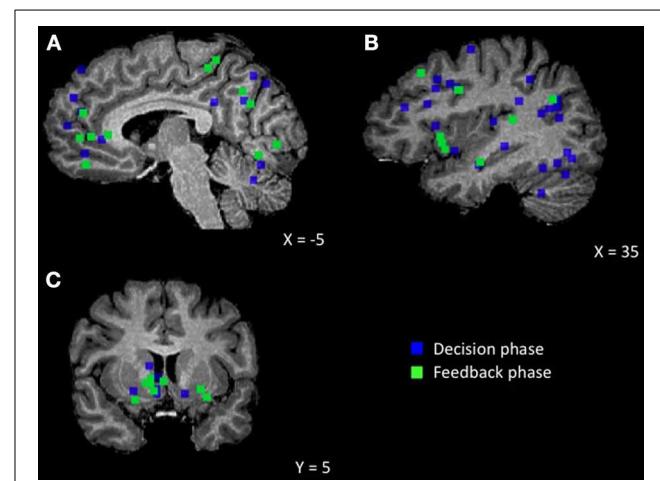


FIGURE 1 | Brain regions showing an effect of human agent compared to non-social control. (A) Medial view displaying MPFC, posterior cingulate/precuneus, cerebellum **(B)** Lateral view displaying STS, TPJ, DLPFC, IPL, insula, fusiform **(C)** Coronal view displaying striatum.

when viewing the person and learning about their past behavior, while also taking the social context into account. Yet discussions of how these predictions are utilized within a decision-making context have eluded social psychology researchers in favor of understanding the processes by which such predictions are made. Below, we discuss these social cognitive processes and how they influence social decision-making in various behavioral economic paradigms involving human and computer agents.

Social decisions are not made within a vacuum; they are made in a social context. A social context involves the actual, imagined, or implied presence of another person—an intentional agent—whose behavior cannot be predicted with certainty. Although humans have developed ways to try to predict what another person will probably do, the other person has the ability to originate their own actions and only they know their true intentions. Therefore, social decision-making is complicated by the uncertainty of the other person's behavior and requires inferences about a person's mental state. Despite these uncertainties, humans are highly motivated to explain and predict others behavior (Heider, 1958). To facilitate this process, humans have developed skills to automatically assess or infer certain types of social information about another person that will guide predictions about their behavior. The primary dimensions of person perception—trait warmth and trait competence—allow for these predictions (Asch, 1946; Rosenberg et al., 1968; Fiske et al., 2007). While trait warmth describes a person's good or bad intentions, trait competence describes the person's ability to carry out those intentions. Research suggests that although these two traits are often assessed together (Fiske et al., 2002), trait warmth carries more weight when forming impressions (Asch, 1946). As such, it is not surprising that the majority of social decision-making studies have capitalized on participants' ability to infer something about warmth-related constructs, including trustworthiness, fairness, and altruism in economic games.

²Rangel et al. (2008) suggests five steps for value based decision-making, including the three listed here as well as a representation stage and an action selection stage. We do not focus on these 2 steps here because they may not be all that different for social and nonsocial decision-making.

But social predictions are not always formed based on trait inferences alone—social category information (e.g., age, race, gender) and physical features (e.g., facial trustworthiness, attractiveness) can guide initial impressions of a person as well (Fiske, 1998; Ito and Urland, 2003; Ito et al., 2004). Stereotypes—schemas about how people belonging to social categories behave—can act as heuristics for predicting a person's behavior based on this category information (Fiske, 1998; Frith and Frith, 2006). However, these predictions can often be misleading because they do not require mental state inferences for the individual person. Despite this, social category information such as gender and race affect social decisions in an economic context (Slonim and Guillen, 2010; Stanley et al., 2011), suggesting this social information is incorporated into the decision-making process when interacting with human agents.

The basis of these social predictions (e.g., social category information, physical features, and trait inferences) are often assessed automatically and efficiently, with only 100 ms of exposure to a person's face leading to accurate assessments (Willis and Todorov, 2006). These initial impressions may be further supported or adjusted based on the person's behavior. People spontaneously attribute traits to a person based on brief, single acts (thin slices) of behavior. When exposure time to a person's behavior is increased from 30 s to 4 to 5 min, predictions about their future behavior are just as accurate as with minimal exposure (Ambady and Rosenthal, 1992). Therefore, these automatic social processes may influence any social decision-making study that has an actual, imagined, or implied presence of another person.

The development of attribution theory (Heider, 1958; Kelley, 1972; Jones, 1979) further suggests that people are highly motivated to predict and explain behavior and are able to do so quite efficiently. Kelley (1972) suggests only three pieces of information—what other people do (consensus), reliability of a behavior across contexts (distinctiveness), and reliability of a behavior across time (consistency)—are needed for participants to form enduring trait inferences and attribute behavior to a person rather than the situation. Specific combinations—low consensus, low distinctiveness, and high consistency—lead participants to attribute behavior to the agent (McArthur, 1972). Interestingly, research shows that this attribution process may be different for social and non-social stimuli. When this paradigm was taken to the scanner, Harris et al. (2005) showed that attributions for human agents rely on a distinct set of brain regions, including MPFC and STS. However, when the agents are anthropomorphized objects, the same combination of statistical information led to attributions (i.e., the same behavior for human and objects) but a different pattern of brain activity resulted (Harris and Fiske, 2008). Specifically attributions for objects did not engage MPFC but rather STS and bilateral amygdala. These studies, in combination with studies showing increased activity in dorsal regions of MPFC for people compared to objects (cars and computers) in an impression formation task (Mitchell et al., 2005) suggest separable brain systems for people and objects and provide a first hint toward what makes social decision-making different.

What does social psychology teach us about social decision-making studies? Participants use a variety of heuristics that allow

them to infer traits and mental states about another person. Whether this is information about their identity (e.g., age, race, gender) or information about their past behavior, participants are constantly trying to make predictions about what other people will do (even outside of a decision-making context). As such, traits provide a concise schema suggesting how a person will behave, allowing for generalizations across contexts when making predictions about behavior. In general, if a person is thought to be trustworthy in one context, people predict that they will be trustworthy in other contexts. Whether actual consistency across contexts exists depends on the psychological viewpoint one takes—personality psychologists would suggest traits are an enduring quality that stays consistent across situations, however, social psychologists stress the importance of the situation and the interaction between person and environment (Lewin, 1951; Ross and Nisbett, 1991).

How does this contribute to our discussion of human and computer agents in an economic game? Do participants use the same brain regions when making predictions about what a human will do vs. what a computer will do? Since each type of agent recruits different brain regions, do social predictions rely on the person perception/social cognition network as we hypothesize above? Below we describe three economic games—the trust game, ultimatum game, and prisoner's dilemma game—often used in the neuroeconomics literature on social decision-making and discuss how social cognition and social psychological theory may be useful when studying these games. We also review research that will help us understand the brain regions underlying these predictions, specifically studies that use non-social agents (e.g., computers) as a control and examine activation during the decision phase when participants are making predictions about what the other agent will do (see **Table 1** for list of studies).

One tool for studying social predictions is the trust game. In a typical trust game scenario, participants have the opportunity to “invest” with or give a sum of money (e.g., \$10) to another person. Alternatively, participants can decide to keep the money for themselves and not invest. If the money is given to the partner, it is multiplied by some factor (e.g., tripled to \$30) and the partner decides whether or not to share the profit with the investor. If the partner shares with the participant, each receives an equal payout (\$15). However, if the partner decides to keep the profit (\$30), the participant receives nothing. Participants must predict what the partner will do in order to maximize their payout. If they predict the partner will not share, the participants should not invest and keep the money for themselves. However, if participants predict the partner will share, the participants should invest with the partner, risking the chance that they will lose the whole amount.

How do participants make these predictions if they have never interacted with their partners before? From a social cognition perspective, spontaneous mental state inferences may guide these predictions, resulting in corresponding activity in social cognition brain regions. In fact, research shows that when making such predictions for human and computer agents in a trust game social cognition brain regions including the prefrontal cortex (PFC) and inferior parietal cortex (IPL) are more active for human compared to computer partners when participants decide to invest (McCabe et al., 2001; Delgado et al., 2005). However,

Table 1 | Summary of studies comparing human and non-social agents.

Phase	Author	Method	Task	Nonsocial comparison	Brain regions associated with an effect of human agent
Decision	McCabe et al., 2001	fMRI	Trust game	Computer	MPFC
Decision	Gallagher et al., 2002	PET	Rock-Paper-Scissors	Computer	pACC
Decision	Singer et al., 2004	fMRI	PDG	Nonintentional human	fusiform gyrus, STS, insula, vSTR, OFC
Decision	De Quervain et al., 2004	fMRI	Punishing defector in trust game	Random device	caudate nucleus
Decision	Rilling et al., 2004a	fMRI	UG and PDG	Computer	DLPFC, STG, fusiform gyrus, precentral gyrus, inferior frontal gyrus, superior frontal gyrus, posterior cingulate, frontal pole, caudate, cerebellum
Decision	Delgado et al., 2005	fMRI	Trust game	Lottery	IPL, insular cortex, lingual gyrus, putamen, inferior occipital gyrus, vSTR, fusiform gyrus
Decision	Knoch et al., 2006	fMRI	UG	Computer	DLPFC
Decision	Krach et al., 2008	fMRI	PDG	Anthropomorphized robot, functional robot, computer	MPFC, TPJ
Decision	Coricelli and Nagel, 2009	fMRI	Beauty contest	Computer	MPFC, rACC, STS, PCC, TPJ
Decision	Burke et al., 2010	fMRI	Purchasing stocks	Chimpanzees	vSTR
Decision	Carter et al., 2012	fMRI	Poker game/bluffing decisions	Computer	TPJ
Decision	Delgado et al., 2008	fMRI	Auction	Lottery controlled by computer	precuneus, inferior parietal lobe
Feedback	Rilling et al., 2002	fMRI	PDG	Computer	paracentral lobule, caudate, postcentral gyrus, medial frontal gyrus, rostral anterior cingulate gyrus, superior temporal gyrus, paracentral lobule
Feedback	Sanfey et al., 2003	fMRI	UG	Computer	bilateral insula
Feedback	Rilling et al., 2004b	fMRI	UG and PDG	Computer	STR, VMPFC
Feedback	Rilling et al., 2004a	fMRI	UG and PDG	Computer and Roulette Wheel	STS, hypothalamus/midbrain/thalamus, superior frontal gyrus, rACC, precuneus, thalamus, hippocampus, putamen
Feedback	Delgado et al., 2005	fMRI	Trust game	Lottery	STR (neutral human)
Feedback	Rilling et al., 2008a	fMRI	PDG	Gamble task	superior temporal gyrus, precentral gyrus, anterior insula, precuneus, lingual gyrus, ACC
Feedback	Delgado et al., 2008	fMRI	Auction	Lottery controlled by computer	STR
Feedback	Phan et al., 2010	fMRI	Trust game	Computer	vSTR

(Continued)

Table 1 | Continued

Phase	Author	Method	Task	Nonsocial comparison	Brain regions associated with an effect of human agent
Feedback	Harlé et al., 2012	fMRI	UG	Computer (between group contrast)	anterior insula, OFC, DLPFC, precentral gyrus, superior temporal pole, vMPFC, lateral prefrontal cortex, putamen, SMA, parahippocampal Area, precuneus, ACC, cerebellum, inferior parietal gyrus

Brain regions associated with an effect of human agent (compared to non-social control) include social cognition brain regions. UG, ultimatum game; PDG, prisoner's dilemma game; MPFC, medial prefrontal cortex; pACC, posterior anterior cingulate cortex; STS, superior temporal sulcus; vSTR, ventral striatum; OFC, orbital frontal cortex; DLPFC, dorsolateral prefrontal cortex; STG, superior temporal gyrus; TPJ, temporal parietal junction; PCC, posterior cingulate cortex; SMA, supplemental motor area.

no differences are observed in activation when participants do not invest, suggesting that investing in the trust game requires inferring the mental states of the partner.

Past behavior may also inform predictions in the trust game. Remember that people form trait inferences from brief single acts of behavior. In a trust game situation, the partner's decision will allow the participant to infer that the partner is trustworthy (or not) from a single exchange. If this behavior is repeated, the partner will build a reputation (a trait inference) for being trustworthy. When relying on reputation to predict the partner's actions, striatal activation shifts from the feedback phase when processing rewards to the decision phase when viewing pictures of previous cooperators, suggesting that participants are making predictions that previous cooperators will again cooperate in the current trial (King-Casas et al., 2005). Therefore, the striatum is also involved in forming social predictions.

Similarly, participants in the ultimatum game interact with human and computer agents that propose different ways of dividing a sum of money (e.g., \$10). While some of these offers are fair (\$5 each party), others are unfair (\$3 for the participant and \$7 for the partner). If the participant decides to accept the offer, the money is divided as proposed. However, if the participant rejects the offer, both parties receive nothing. In an economic sense, any non-zero offer should be accepted in order to maximize payout, especially if partners are not repeated throughout the experiment (one-shot games). However, research suggests that unfair offers are rejected more often when the partner is a human agent than computer agent. Why does the identity of the partner affect decisions if the same economic outcome would result? Perhaps, related to our discussion of flexibility above, participants know that humans respond to the environment and make adaptive decisions. If they see that their unfair offers are being rejected, the participant may predict that the human partner will change their behavior, offering more fair offers. However, a computer may be predicted to propose the same offer regardless of how the participant responds, in which case it would be advantageous to accept any non-zero offer because the participant does not anticipate the computer would respond to his or her rejection of the offers. Rejection may also represent a form of punishment of the partner. If the participant receives a low offer, this suggests that the partner has a negative impression of the participant or is simply a morally bad person (unfair, selfish). Punishment in this light is

action against such mental states. However, since computers do not possess mental states, there is no reason to punish them for similar unfair offers.

Research shows that when deciding whether to accept or reject offers proposed by human and computer agents, participants show higher skin conductance responses to unfair offers made by human compared to computer agents (Van't Wout et al., 2006), suggesting increased emotional arousal. The use of repetitive transcranial magnetic stimulation (rTMS) shows disruption of the right dorsolateral prefrontal cortex (DLPFC) leads to higher acceptance rates of unfair offers from human but not computer agents (Knoch et al., 2006). The authors of this study highlight the role of DLPFC in executive control and suggest this region is essential for overriding selfish impulses in order to reject unfair offers. When this region is disrupted, participants are more likely to act selfishly and are less able to resist the economic temptation of accepting any non-zero offer. Although the role of DLPFC in executive control is not debated, a more social psychological explanation may be useful in understanding this behavior as well. Impression management is believed to be part of executive control function (Prabhakaran and Gray, 2012). Therefore, we may ask if DLPFC is involved in overriding selfish impulses specifically or whether concerns about impression management may also be affected by the DLPFC's role in executive control. Accepting and rejecting offers in the ultimatum game communicates something to the partner about the participant—whether or not they will accept unfair treatment. In other words, the participant's behavior allows the partner to (presumably) form an impression of them. In order to manage this impression, participants may reject unfair offers as a way to communicate that he or she will not stand for being treated unfairly. Therefore, perhaps when DLPFC is disrupted with rTMS, impression management concerns are reduced and unfair offers are more often accepted. Concerns about forming a good reputation are also affected by rTMS to right DLPFC in the trust game (Knoch et al., 2009), further suggesting this region may be involved in impression management.

The prisoner's dilemma game (PDG) is another economic game exemplifying the role of predictions in social decision-making. In this game, participants must decide whether to cooperate with a partner for a mediocre reward (e.g., \$5 each), or defect in order to receive a better reward at the expense of the partner (e.g., \$10 for the participant, \$0 for the partner).

However, risk is introduced into the game because if the partner also defects, both players end up with the worst possible outcome (e.g., \$0). In this case it is important for the participant to predict what the partner will do because the payout structure that both parties receive depends on what each chooses.

When participants believe they are playing with human rather than computer agents, imaging results show greater activation in regions involved in social cognition, including right posterior STS, PCC, DLPFC, fusiform gyrus, frontal pole, along with decision-making regions like the caudate (Rilling et al., 2004a). Time-course data show specifically within posterior STS and PCC there is an increase in activation in response to the human partner's face that remains elevated until the outcome is revealed. This increase in activity in social cognition brain regions to human partners is further supported by a study examining PDG decisions to agents varying in degree of human-likeness. Participants that played the PDG with a human, anthropomorphized robot (human-like shape with human-like hands), functional robot (machine-like shape with machine-like hands), and computer showed a linear increase in MPFC and right TPJ activity as human-likeness increased (Krach et al., 2008).

In addition to the agent's perceived physical likeness to a human, it seems as though the intentionality of the human agents is essential for activating social cognition regions. In a study that manipulated whether human agents were able to decide freely in the PDG (intentional) vs. following a predetermined response sequence (unintentional), Singer et al. (2004) observed increased activation of posterior STS, bilateral fusiform gyrus, bilateral insula, right and left lateral OFC, and ventral striatum for cooperating intentional humans. Therefore, it is not that all humans activate social cognition regions in the PDG, but specifically intentional human agents. Together these studies suggest activity in social cognition brain regions track whether the partner is a social agent and may influence social decisions.

Although these economic games are most often used to study social decision-making, other games also suggest that social cognition brain regions are essential for predicting the actions of others. For instance, when playing a game of Rock-Paper-Scissors with either a human or computer counterpart, Gallagher et al. (2002) observed bilateral activation in pACC for human compared to computer partners. More recently, the TPJ has been identified as providing unique information about decisions involving social agents. Participants playing a poker game with human and computer agents had to predict whether the agent was bluffing. Using MVPA and a social bias measure, Carter et al. (2012) showed that TPJ contains unique signals used for predicting the participant's decision specifically for socially relevant agents but not for computer agents. And lastly, research suggests there are individual differences in the extent to which people use social cognition in a decision-making context. In the beauty contest game, participants must choose a number between 0 and 100 with the aim of choosing a number that is closest to 2/3 times the average of all the numbers chosen by different opponents. When playing this game with human and computer opponents, Coricelli and Nagel (2009) found that human opponents activated regions involved in social cognition, including MPFC, rostral ACC, STS, PCC, and bilateral TPJ. The researchers then examined individual

differences in participants' ability to think about others' mental states. While low-level reasoners do not take into account the mental states of others when guessing, high-level reasoners think about the fact that others are thinking about the mental states of others and try to guess accordingly. Interestingly including this individual difference measure in the analysis showed that activity in MPFC was only significant for high-level reasoners.

Together, across different social decision-making paradigms, there seems to be increasing evidence that human and computer agents engage different brain regions when making predictions. Specifically, making predictions about human agents engages brain regions implicated in the social cognition network, including MPFC, STS, TPJ, along with decision-making regions like the striatum. Next we ask whether these social decision-making paradigms engage different brain circuitry when processing feedback from human and computer agents.

SOCIAL FEEDBACK

While many studies have suggested that social predictions rely on the social cognition brain network, other social decision-making studies have looked at how the outcome of social decision-making, or social feedback, affects traditional decision-making brain regions involved in reward processing and valuation. Initial attempts to study the uniqueness of social decision-making include examining whether social and non-social rewards are processed in the same areas of the brain, and how economic decisions are made in the context of social constructs including trustworthiness, fairness, altruism, and the like. Using behavioral economic games described above (e.g., trust game, ultimatum game, etc.) researchers have examined the influence of positive and negative feedback on social decisions. Below, we review the results of such studies in an attempt to continue the comparison between human and computer agents in social decision-making.

Social feedback often allows people to infer something about another person as well as receive information about the impression others have formed of them. In the context of receiving direct social feedback about what other people think, research suggests that being labeled trustworthy activates the striatum in much the same way as receiving monetary rewards (Izuma et al., 2008). This concept of trust is important when making decisions in a social context because it affects existing social interactions as well as whether others will interact with you. In the economic trust game described above, feedback about whether or not the partner returns an investment allows for trait inferences about the partner based on thin slices of behavior that may guide future predictions.

When participants play the trust game with another human, reward related regions such as the caudate nucleus are active (King-Casas et al., 2005). With repeated exposure to the partner's behavior, participants form a reputation (an inferred trait) for the partner as being trustworthy or not. When these partners are human and computer agents, participants differentiate cooperating from non-cooperating humans, investing most often with humans that returned the investment, an average amount with a neutral human, and least often with humans that did not return the investment. Investments for the computer agent were similar to the neutral human. Reflecting this pattern of behavior, brain activity within the left and right ventral striatum reveals

increased activity to cooperating compared to non-cooperating humans, but activity to computers looks similar to neutral human partners (Phan et al., 2010). These results suggest that if a human agent provides no informative information that allows for a trait inference (a neutral partner is neither good or bad), behavior and brain activity may be similar to that of a computer agent. Similar results are observed when reading descriptions of hypothetical partners' past moral behaviors. When playing the trust game with a neutral investment partner (neither good or bad moral character) activity within the striatum for positive and negative feedback looks similar to when receiving such feedback about a non-social lottery outcome (Delgado et al., 2005). However, when the human agent is associated with a specific moral character, striatal activity for positive and negative feedback look the same, demonstrating that prior social information can bias feedback mechanisms in the brain, but only when the social information is informative about one's traits.

In the trust game, the outcome phase has a clear start and end—participants make a decision to invest (share) with a partner and then receive feedback in the same trial about whether the investment was returned by the partner. However, in the ultimatum game, the outcome phase is less clear—participants already know the outcome of the social interaction when they decide whether to accept or reject the offer made by the agent. However, this does not make the outcome of the social interaction irrelevant. In repeated ultimatum games (when participants play multiple trials with the same partner), feedback about the participant's decision comes on the next trial when the partner proposes the next division of money. For example, if a participant rejects an unfair offer, feedback about whether that rejection was effective in influencing the partner's next proposal comes on the next trial. In other words, offers can be thought of as feedback within the context of this game. However, researchers often use single-shot ultimatum games to avoid effects of repeated interaction just described. In this case, the offers proposed by the partner allow the participant to infer traits about the partner, and their decision still communicates something to the partner, prompting participants to think about impression management.

How then do participants respond to offers made by human and computer agents in the context of the ultimatum game? Research suggests that unfair offers made by human agents activate bilateral anterior insula to a greater extent than the same unfair offers made by computer agents, suggesting that there is something about being mistreated specifically by human agents that leads to higher rejection rates (Sanfey et al., 2003). Additionally it seems as though the balance of activity in two regions—anterior insula and DLPFC—predicts whether offers are accepted or rejected. Unfair offers that are subsequently rejected have greater anterior insula than DLPFC activation, whereas accepted offers exhibit greater DLPFC than anterior insula. Similarly, when viewing a human partner's offer, social cognition and decision-making regions including STS, hypothalamus/midbrain, right superior frontal gyrus (BA8), dorsal MPFC (BA 9, 32), precuneus, and putamen are active (Rilling et al., 2004a). More recent investigations of unfair offers suggest the identity of the agent (human or computer) determines whether mood has an effect on activity in bilateral anterior insula (Harlé

et al., 2012). Specifically, sad compared to neutral participants elicited activity in anterior insula and ACC as well as diminished sensitivity in ventral striatum when viewing unfair offers from human agents but there were no such differences for offers made by computer agents. These differences in brain activity for human and computer agents further highlight that social decision-making (compared to non-social) relies on different neural processing.

Unlike the ultimatum game, the prisoner's dilemma game is similar to the trust game, because the participant and the partner must make a decision before finding out the outcome of both parties' decisions. This outcome period lets the participant know whether their predictions about the partner were correct. When participants played the prisoner's dilemma game in the scanner, Rilling et al. (2002) observed different patterns of brain activation during outcome depending on whether the partner was a human or computer agent. Specifically, both human and computer agents activated ventromedial/orbital frontal cortex (BA 11) after a mutually cooperative outcome (both the partner and participant decided to cooperate). However, mutual cooperation with human partners additionally activated rostral anterior cingulate and anteroventral striatum. A few years later, researchers investigated whether these different activations were limited to when partners cooperate. Comparing social to non-social loss (human partners do not cooperate and losing a monetary gamble), Rilling et al. (2008a) observed higher activation in superior temporal gyrus (BA 22), precentral gyrus, anterior insula, precuneus, lingual gyrus, and anterior cingulate for the human agent. This analysis highlights the importance of human agents' perceived intent in the prisoner's dilemma game, as it controls for differences in monetary payoff, frequency, and emotional valence that may have confounded previous comparisons of cooperation and defection. These studies suggest processing outcomes from human and computer agents is different. Specifically, human agents engage social cognition brain regions, perhaps because outcomes lead to spontaneous trait inferences for humans and not computers. This idea is consistent with social neuroscience research showing different activity when attributing behavior to people and objects (Harris et al., 2005; Harris and Fiske, 2008).

In another study, participants played a time estimation task in which a human or computer agent delivered trial-by-trial feedback (juice reward or bitter quinine). Some brain regions, including ventral striatum and paracingulate cortex (PACC) responded more to positive vs. negative feedback irrespective of whether the agent was a human or computer (Van den Bos et al., 2007). Other brain regions, particularly bilateral temporal pole, responded more to feedback from human than computer agents, regardless of feedback valence. However, the combination of type of agent and feedback valence seems to be important within the regions of anterior VMPFC and subgenual cingulate. Interestingly this study is one of the few comparing human and computer feedback that is relevant to the competence rather than warmth domain but delivers the same take home message—some brain regions like the striatum and prefrontal cortex respond to social and non-social stimuli, but others like social cognition regions are engaged specifically to the human agent. Why are social cognition regions engaged if feedback was dependent on the participant's

performance in the task and not the agents' decisions (i.e., delivered feedback did not allow for a trait inference about the agent)? It may be that participants were concerned about the impression the human agent formed of them (i.e., participants know their behavior allows for trait inferences about them in the same way they form trait inferences about others), but these concerns were not relevant for the computer agent because computers do not form impressions.

Another study examining the effects of competing against a human or computer in an auction suggests that differences in brain activity during outcome depend on both the type of agent and the context of the outcome (Delgado et al., 2008). Participants were told that they would be bidding in an auction against another human or playing a lottery game against a computer and had the opportunity to win money or points at the end of the experiment. The points contributed to the participant's standing at the end of the experiment in which all participants would be compared. In other words, the points represented a social reward, allowing participants to gain status when comparing themselves to other participants in the study. In both cases the goal was to choose a number higher than that chosen by the other agent. When the outcome of the bidding was revealed, the authors observed differential activity for the social and lottery trials. Specifically, losing the auction in the social condition reduced striatal activity relative to baseline and the lottery game. The authors suggest that one possible explanation for overbidding in auctions is the fear of losing a social competition, which motivates bids that are too high, independent from pure loss aversion. These differences for social and non-social loss highlight again that although the same brain regions are active, the social context modulates activity within decision-making regions.

But should we be surprised that social loss seems more salient to participants in a social competition such as the one created by the experimenters? Specifically, the experimenters told participants that final results about the participant's standing in relation to other participants would anonymously be released at the end of the study in a list of "Top 10 players." Even though there was no risk of identifying a particular participant, social concerns about impression management may have still been active. Being listed as one of the top players allows the trait inference of being very competent in the auction, a desirable trait to almost anyone. Therefore, participants may have believed that negative feedback (losing the auction trials) would lead people to infer that they were inferior or incompetent compared to other players. On the other hand, losses on the lottery trials were simply relevant to the participants and not their social standing.

Converging evidence suggests that common brain regions, particularly the striatum and VMPFC, are engaged when viewing outcomes from human and computer agents. However, the activity in these regions seems to be modulated by the social context. In addition to these decision-making regions, the ultimatum game and prisoner's dilemma game also activate regions involved in social cognition, including STS, precuneus, and TPJ. Should it be surprising that social cognition regions are also active during outcomes? Social psychology demonstrates that people infer traits from others' behavior. The outcome of a social interaction allows participants to infer these traits, and what perhaps is even more

interesting is that these trait inferences are formed in single-shot games where participants do not interact with the partner again. Essentially, trait inferences in this context are superfluous because the participant will not be interacting with the partner again so there is no need to infer traits that allow for predictions. Yet these social cognition regions are still engaged.

SOCIAL LEARNING

So far we have seen that social cognition informs predictions made in social decision-making studies when interacting with human but not (or to a lesser extent) when interacting with computer agents. Social rewards, including being labeled trustworthy by another person (Izuma et al., 2008), gaining social approval by donating money in the presence of others (Izuma et al., 2010), and viewing smiling faces (Lin et al., 2012) engage brain regions that are common to receiving non-social rewards, such as money. However, when receiving feedback from social and non-social agents, though common brain regions including the striatum are engaged, the type of agent may modulate activity in these regions. Moreover, feedback from a social interaction also engages regions of the social cognition network. Next, we examine differences in social decision-making during the updating or learning process.

Research examining learning in a non-social context has highlighted the role of prediction error signals in learning to predict outcomes. In a now classic study, recordings from dopamine neurons show that primates learn to predict a juice reward, shifting the firing of dopamine neurons to the cue rather than reward. When an expected reward is not received, dopamine neurons decrease their firing (Schultz et al., 1997). Similar prediction error signals have been observed to social stimuli in both an attribution task (Harris and Fiske, 2010) as well as in decision-making contexts (King-Casas et al., 2005; Rilling et al., 2008b for review). In recent years, it has therefore been suggested that social learning is akin to basic reinforcement learning (i.e., social learning is similar to non-social learning). When interacting with peers, ventral striatum and OFC seem to track predictions about whether a social agent will give positive social feedback and ACC correlates with modulation of expected value associated with the agents (Jones et al., 2011). It has also been proposed that social information may be acquired using the same associative processes assumed to underlie reward-based learning, but in separate regions of the ACC (Behrens et al., 2008). These signals are believed to combine within MPFC when making a decision, consistent with the idea of a common valuation system (which combines social and non-social) within the brain (Montague and Berns, 2002). In fact, value signals for both social and monetary rewards have been found to rely on MPFC (Smith et al., 2010; Lin et al., 2012) and activity in this region also correlates with the subjective value of donating money to charity (Hare et al., 2010).

However, social learning does not inherently appear to be just another type of reinforcement learning. Social decisions often contradict economic models that attempt to predict social behavior, suggesting that simple reinforcement learning models by themselves are not sufficient to explain complex social behavior (Lee et al., 2005). Research shows that reward and value signals are modulated by the social context. For instance, reward related signals in the striatum are affected by prior social information

about an investment partner (Delgado et al., 2005) as well as when sharing rewards with a friend vs. a computer (Fareri et al., 2012). Additionally, research shows that social norms can influence the value assigned to social stimuli, specifically modulating activity in nucleus accumbens and OFC (Zaki et al., 2011). Interestingly, functional connectivity analyses show that value signals in MPFC may rely on information from person perception brain regions like the anterior insula and posterior STS (Hare et al., 2010). Studies investigating how person perception brain regions affect social learning suggest that specific types of social information (warmth vs. competence) affect social learning—whereas information about a person's warmth hinders learning, information about a person's competence seems to produce similar learning rates as when interacting with computer agents (Lee and Harris, under review).

Should we be surprised by findings that social stimuli affect learning and the updating process? Social psychology suggests the answer to this question is no. Behaviorally, people have a number of biases that may affect the way information is processed and incorporated into decision-making processes. Tversky and Kahneman (1974) were perhaps the first to point out these biases and heuristics that may be used in a social decision-making context. For instance, people use probability information to judge how representative a person is of a specific category (representativeness heuristic), and recent events to assess how likely it is that something will occur (availability heuristic). When asked to give an estimate of some quantity, being given a reference point (an anchor) affects the resulting estimates. These heuristics can be applied to a social decision-making context as well. For instance when playing the trust game, participants may use initial impressions formed about the person (based on a representative heuristic about what trustworthy people look like) as an anchor that affects whether or not they invest with the partner on subsequent trials. In addition to this bias, it is harder to verify cognitions about people than objects, making it harder to accurately infer the traits of a person compared to an object (Fiske and Taylor, 2013).

In addition to the heuristics described above, people also possess a number of biases that affect how they interpret information. First, people look for information that is consistent with a preexisting belief. This confirmatory bias is evident in the stereotype literature, which demonstrates that people interpret ambiguous information as consistent with or as a confirmation of a stereotype about a person (Bodenhausen, 1988). This bias is relevant to the economic games employed in social decision-making studies because partners often provide probabilistic (sometimes ambiguous) feedback. Interpretation of this feedback may be influenced by prior beliefs (Delgado et al., 2005). Second, people often exhibit illusionary correlations—that is they see a relationship between two things when one does not exist (Hamilton and Gifford, 1976)—and are more likely to attribute a person's behavior to the person rather than to some situational factor (Jones and Davis, 1965; Jones and Harris, 1967; Ross, 1977; Nisbett and Ross, 1980). This again leads participants in social decision-making studies more likely to interpret a partner's decision as a signal of some underlying mental state or trait attribute rather

than positive or negative feedback in a purely reward processing sense.

How then can we reconcile these two different literatures, one stating that social learning is similar to reinforcement learning, and another stating that social learning includes a number of biases? In more practical terms, we know that impressions of a person can guide decision-making. Previous studies have shown that facial trustworthiness affects investment amounts in the trust game (Van't Wout and Sanfey, 2008). However, first impressions are not the only influence on social decisions—if someone is perceived as trustworthy that does not make their subsequent behavior irrelevant. Other research has shown the importance of prior behavior on trust decisions (Delgado et al., 2005; King-Casas et al., 2005). To study how the combination of impressions and behavior affect social decision-making, Chang et al. (2010) used mathematical models based on reinforcement learning to test specific hypotheses about how these two types of information guide social decisions in a repeated trust game. Specifically, the authors tested three models that suggest different ways of processing information and investigate whether reinforcement learning or social biases influence decision-making. First, an Initialization model assumes that initial impressions (implicit trustworthiness judgments) influence decision-making at the beginning of the trust game, but eventually participants learn to rely on the player's actual behavior. A Confirmation Bias model assumes that initial impressions of trustworthiness affect the way feedback is processed, the impression is updated throughout the study, and learning is biased in the direction of the initial impression. The third, Dynamic Belief model, assumes that initial impressions are continuously updated based on the participant's experiences in the trust game and these beliefs then influence learning. In this model, equal emphasis is placed on the initial judgment and the participant's experience. That is, initial trustworthiness is simultaneously influencing learning and being updated by experience. Of the three models, the Dynamic Belief model fit the data the best, suggesting that both social cognition processes (initial impressions) and decision-making processes (feedback processing) affect social learning in the trust game.

More recent social decision-making studies have investigated how social processes affect learning. Researchers have proposed different strategies participants may use when learning to predict what their partner will do. One such strategy is learning to simulate other people's decisions and update those simulations once the other's choice is revealed. This process engages different regions of prefrontal cortex involved in valuation and prediction error (Suzuki et al., 2012). Another strategy is to account for the influence one's decisions have on the partner's decisions and decide accordingly. This strategy requires predicting how much influence one has on the partner and updating that influence signal when observing the partner's decision. Computational modeling suggests MPFC tracks the predicted reward given the amount of expected influence the participant's choices have on the partner, and STS activity is responsible for updating the influence signal (Hampton et al., 2008). Although these studies do not provide direct comparisons to non-social controls, they provide exciting insight into how social cognition processes affect social learning.

CONCLUSION

Is social decision-making unique? How does it differ from non-social decision-making? The answers to these questions have been of interest to researchers in a variety of fields including social psychology and behavioral economics. Combining these literatures can help us understand the answers to these questions. Economists originally believed that social decision-making was not different from non-social decision-making and tried to model social decisions with traditional economic models. However, after the influential paper by Tversky and Kahneman (1974) demonstrating heuristics and biases affecting decision-making, it became apparent that the decision-making process is not as rational as we may have originally thought. Psychologists have long believed that social cognition is important for predicting the actions of others and that humans are different from objects in some very important ways. More recently, brain-imaging studies have highlighted these differences, with a network of brain regions responding to social stimuli and social cognitive processes that presumably affect social decision-making. Investigations of social decisions have also highlighted the effects of social information on decision-making processes within brain regions like the striatum and MPFC. Although both social and non-social agents engage these brain regions, the social context modulates this activity. The use of mathematical models suggests that both social neuroscience and neuroeconomics studies have each been tapping into different processes. Initial impressions allow for predictions that guide decision-making. These impressions then interact with feedback processing and affect how predictions are updated.

In economics, behavioral game theorists recognize that people's beliefs about others matter when modeling social decisions. The models assume that players strategically choose options that maximize utility, and evaluations of payoff options often include social factors beyond pure economic payout (Camerer, 2009). These social factors may include other-regarding preferences, indicating that people care about the well-being of other players (Fehr, 2009). Whether decisions are made in order to increase the well-being of others or manage the impression formed of oneself, mental state inferences are still relevant. For instance, one may assess well-being by inferring the mental state of the person. Similarly, the extent to which one infers the mental state of a person may influence the extent to which other-regarding preferences influence decisions (e.g., do people show other-regarding preferences for traditionally dehumanized targets?).

Humans evolved in a social context in which interacting with other people was essential for survival. As such, these social cognitive processes have been evolutionarily preserved and continue to affect our decision-making in a social context. The fact that human agents engage different brain regions than computer agents should perhaps not be all that surprising. The social brain did not evolve interacting with computers or other types of machines. Therefore, we see differences not only in behavior (most of the time) but also differences in brain activity for these two inherently different types of agents. Here we have highlighted that these differences lie in engagement of the social cognition/person perception brain regions for human agents. But the underlying mechanisms—the social processes that engage these brain regions and how they interact with decision-making

processes—are still being investigated. Social psychological theory can help answer these questions by providing a theoretical background for why human and computers differ in the first place (e.g., mental state inferences, impression management, etc.). Keeping this fact in mind will provide future research on social decision-making with the most informed and cohesive theories.

Finally, decisions are made in a social context everyday. Whether deciding to do a favor for a friend or close a deal with a potential business partner, decisions have consequences that lead to significant rewards and punishments such as a better relationship with the friend or a poor business transaction. Therefore, it is important to understand how decisions are influenced by the presence or absence of others and how we incorporate social information into our decision-making process. Here we have highlighted differences arising when interacting with human and computer agents and use social psychological theory to provide some explanation for why these differences arise. It is important to point out these differences in social and non-social decision-making because interactions with computers and other machines are becoming more widespread. Businesses often try to find ways to simplify transactions, often replacing human agents with automated computers. However, the decisions made with these different types of agents may affect businesses in unanticipated ways. Financial decisions (e.g., buying and selling stock) are increasingly made through the use of online computers, whereas previously investors had to interact with stockbrokers in an investment firm. Similarly people are able to bid in online auctions for a desired item rather than sitting in a room full of people holding numbered paddles. The decisions to buy and sell stock or possibly overbid in an online auction may be influenced by these different agents, as evidenced by the research described above.

REFERENCES

- Alexander, W. H., and Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 10, 1338–1346. doi: 10.1038/nn.2921
- Ambady, N., and Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychol. Bull.* 111, 256–274. doi: 10.1037/0033-2909.111.2.256
- Amadio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277. doi: 10.1038/nrn1884
- Asch, S. E. (1946). Forming impressions of personality. *J. Abnorm. Soc. Psychol.* 42, 258–290. doi: 10.1037/h0055756
- Behrens, T. E., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. (2008). Associative learning of social value. *Nature* 456, 245–249. doi: 10.1038/nature07538
- Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: testing process models of stereotype use. *J. Pers. Soc. Psychol.* 55, 726. doi: 10.1037/0022-3514.55.5.726
- Burke, C. J., Tobler, P. N., Schultz, W., and Baddeley, M. (2010). Striatal BOLD response reflects the impact of herd information on financial decisions. *Front. Hum. Neurosci.* 4:48. doi: 10.3389/fnhum.2010.00048
- Camerer, C. F. (2009). “Behavioral game theory and the neural basis of strategic choice,” in *Neuroeconomics: Decision-Making and the Brain*, eds P. W. Glimcher, E. Fehr, A. Rangel, C. Camerer, and R. A. Poldrack (London: Academic Press), 193–206.
- Carter, R. M., Bowling, D. L., Reeck, C., and Huettel, S. A. (2012). A distinct role of the temporal-parietal junction in predicting socially guided decisions. *Science* 337, 109–111. doi: 10.1126/science.1219681
- Chang, L. J., Doll, B. B., van't Wout, M., Frank, M. J., and Sanfey, A. G. (2010). Seeing is believing: trustworthiness as a dynamic belief. *Cogn. Psychol.* 61, 87–105. doi: 10.1016/B978-0-12-374176-9.00013-0

- Chib, V. S., Rangel, A., Shimojo, S., and O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29, 12315–12320. doi: 10.1523/JNEUROSCI.2575-09.2009
- Coricelli, G., and Nagel, R. (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9163–9168. doi: 10.1073/pnas.0807721106
- Delgado, M. R., Frank, R. H., and Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618. doi: 10.1038/nn1575
- Delgado, M. R., Schotter, A., Ozbay, E. Y., and Phelps, E. A. (2008). Understanding overbidding: using the neural circuitry of reward to design economic auctions. *Science* 321, 1849–1852. doi: 10.1126/science.1158860
- Delgado, M. R. (2007). Reward-related responses in the human striatum. *Ann. N.Y. Acad. Sci.* 1104, 70–88. doi: 10.1196/annals.1390.002
- De Quervain, D. J. F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., et al. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258. doi: 10.1126/science.1100735
- Engelmann, J. B., and Hein, G. (2013). Contextual and social influences on valuation and choice. *Prog. Brain Res.* 202, 215–237. doi: 10.1016/B978-0-444-62604-2.00013-7
- Fareri, D. S., Niznikiewicz, M. A., Lee, V. K., and Delgado, M. R. (2012). Social network modulation of reward-related signals. *J. Neurosci.* 32, 9045–9052. doi: 10.1523/JNEUROSCI.0610-12.2012
- Fehr, E. (2009). "Social preferences and the brain," in *Neuroeconomics: Decision-Making and the Brain*, eds P. W. Glimcher, E. Fehr, A. Rangel, C. Camerer, and R. A. Poldrack (London: Academic Press), 215–232.
- Fiske, S. T. (1998). "Stereotypes, prejudice, and discrimination," in *Handbook of Social Psychology, 4th Edn*, Vol. 2, eds D. T. Gilbert, S. T. Fiske, and G. Lindzey (New York, NY: McGraw-Hill), 357–411.
- Fiske, S. T., Cuddy, A. J., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Soc. Psychol.* 82, 878–902. doi: 10.1037/0022-3514.82.6.878
- Fiske, S. T., Cuddy, A. J. C., and Glick, P. (2007). Universal dimensions of social perception: warmth and competence. *Trends Cogn. Sci.* 11, 77–83. doi: 10.1016/j.tics.2006.11.005
- Fiske, S. T., and Taylor, S. E. (2013). *Social Cognition: From Brains to Culture* (2/e). London: Sage.
- Frith, C. D., and Frith, U. (2006). How we predict what other people are going to do. *Brain Res.* 1079, 36–46. doi: 10.1016/j.brainres.2005.12.126
- Frith, U., and Frith, C. (2001). The biological basis of social interaction. *Curr. Dir. Psychol. Sci.* 10, 151–155. doi: 10.1111/1467-8721.00137
- Gallagher, H. L., Jack, A. I., Roepstorff, A., and Firth, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage* 16, 814–821. doi: 10.1006/nimg.2002.1117
- Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn. Sci.* 8, 396–403. doi: 10.1016/j.tics.2004.07.002
- Hamilton, D. L., and Gifford, R. K. (1976). Illusory correlation in interpersonal perception: a cognitive basis of stereotypic judgments. *J. Exp. Soc. Psychol.* 12, 392–407. doi: 10.1016/S0022-1031(76)80006-6
- Hampton, A. N., Bossaerts, P., and O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6741–6746. doi: 10.1073/pnas.0711099105
- Hare, T. A., Camerer, C. F., Knoepfle, D. T., and Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* 30, 583–590. doi: 10.1523/JNEUROSCI.4089-09.2010
- Harlé, K. M., Chang, L. J., van't Wout, M., and Sanfey, A. G. (2012). The neural mechanisms of affect infusion in social economic decision-making: a mediating role of the anterior insula. *Neuroimage* 61, 32–40. doi: 10.1016/j.neuroimage.2012.02.027
- Harris, L. T., and Fiske, S. T. (2008). Brooms in Fantasia: neural correlates of anthropomorphizing objects. *Soc. Cogn.* 26, 209–222. doi: 10.1521/soco.2008.26.2.210
- Harris, L. T., and Fiske, S. T. (2010). Neural regions that underlie reinforcement learning are also active for social expectancy violations. *Soc. Neurosci.* 5, 76–91. doi: 10.1080/17470910903135825
- Harris, L. T., Todorov, A., and Fiske, S. T. (2005). Attributions on the brain: neuroimaging dispositional inferences beyond theory of mental state. *Neuroimage* 28, 763–769. doi: 10.1016/j.neuroimage.2005.05.021
- Haxby, J. V., Gobbini, M. I., and Montgomery, K. (2004). "Spatial and temporal distribution of face and object representations in the human brain," in *The Cognitive Neurosciences*, ed M. Gazzaniga (Cambridge, MA: MIT Press), 889–904.
- Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York, NY: Wiley. doi: 10.1037/10628-000
- Ito, T. A., Thompson, E., and Cacioppo, J. T. (2004). Tracking the time-course of social perception: the effect of racial cues on event-related brain potentials. *Pers. Soc. Psychol. Bull.* 30, 1267–1280. doi: 10.1177/0146167204264335
- Ito, T. A., and Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *J. Pers. Soc. Psychol.* 85, 616–626. doi: 10.1037/0022-3514.85.4.616
- Izuma, K., Daisuke, S., and Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron* 58, 284–294. doi: 10.1016/j.neuron.2008.03.020
- Izuma, K., Saito, D. N., and Sadato, N. (2010). Processing of the incentive for social approval in the ventral striatum during charitable donation. *J. Cogn. Neurosci.* 22, 621–631. doi: 10.1162/jocn.2009.21228
- Jones, E. E. (1979). The rocky road from acts to dispositions. *Am. Psychol.* 34, 107. doi: 10.1037/0003-066X.34.2.107
- Jones, E. E., and Davis, K. E. (1965). A theory of correspondent inferences: from acts to dispositions. *Adv. Exp. Soc. Psychol.* 2, 219–266. doi: 10.1016/S0065-2601(08)60107-0
- Jones, E. E., and Harris, V. A. (1967). The attribution of attitudes. *J. Exp. Soc. Psychol.* 3, 1–24. doi: 10.1016/0022-1031(67)90034-0
- Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., et al. (2011). Behavioral and neural properties of social reinforcement learning. *J. Neurosci.* 31, 13039–13045. doi: 10.1523/JNEUROSCI.2972-11.2011
- Kelley, H. H. (1972). "Attribution in social interaction," in *Attribution: Perceiving the Cause of Behaviour*, eds E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, and B. Weiner (Hillsdale, NJ: Lawrence Erlbaum & Associates Inc), 1–26.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Monatgue, P. R. (2005). Getting to know you: reputation and trust in a two person economic exchange. *Science* 308, 78–83. doi: 10.1126/science.1108062
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., and Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832. doi: 10.1126/science.1129156
- Knoch, D., Schneider, F., Schunk, D., Hohmann, M., and Fehr, E. (2009). Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20895–20899. doi: 10.1073/pnas.0911619106
- Knutson, B., and Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* 18, 411–417. doi: 10.1097/01.wco.0000173463.24758.f6
- Krach, S., Hegel, F., Wrede, B., Sagerer, G., Binkofski, F., and Kircher, T. (2008). Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLoS ONE* 3:e2597. doi: 10.1371/journal.pone.0002597
- Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404–409. doi: 10.1038/nn2065
- Lee, D., McGreevy, B. P., and Barraclough, D. J. (2005). Learning and decision making in monkeys during a rock–paper–scissors game. *Cogn. Brain Res.* 25, 416–430. doi: 10.1016/j.cogbrainres.2005.07.003
- Lewin, K. (1951). *Field Theory in Social Science*. New York, NY: Harper and Brothers.
- Lin, A., Adolphs, R., and Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* 7, 274–281. doi: 10.1093/scan/nsr006
- McArthur, L. A. (1972). The how and what of why: some determinants and consequences of causal attribution. *J. Pers. Soc. Psychol.* 22, 171–193. doi: 10.1037/h0032602

- McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11832–11835. doi: 10.1073/pnas.211415698
- Mitchell, J. P., Neil Macrae, C., and Banaji, M. R. (2005). Forming impressions of people versus inanimate objects: social-cognitive processing in the medial pre-frontal cortex. *Neuroimage* 26, 251–257. doi: 10.1016/j.neuroimage.2005.01.031
- Montague, R. P., and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284. doi: 10.1016/S0896-6273(02)00974-1
- Nisbett, R. E., and Ross, L. (1980). *Human Inference: Strategies and Shortcomings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Phan, K. L., Sripada, C. S., Angstadt, M., and McCabe, K. (2010). Reputation for reciprocity engages the brain reward center. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13099–13104. doi: 10.1073/pnas.1008137107
- Prabhakaran, R., and Gray, J. R. (2012). The pervasive nature of unconscious social information processing in executive control. *Front. Hum. Neurosci.* 6:105. doi: 10.3389/fnhum.2012.00105
- Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. doi: 10.1038/nrn2357
- Rilling, J. K., Goldsmith, D. R., Glenn, A. L., Jairam, M. R., Elfenbein, H. A., Dagenais, J. E., et al. (2008a). The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia* 46, 1256–1266. doi: 10.1016/j.neuropsychologia.2007.11.033
- Rilling, J. K., King-Casas, B., and Sanfey, A. G. (2008b). The neurobiology of social decision-making. *Curr. Opin. Neurobiol.* 18, 159–165. doi: 10.1016/j.conb.2008.06.003
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., and Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405. doi: 10.1016/S0896-6273(02)00755-9
- Rilling, J. K., and Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annu. Rev. Psychol.* 62, 23–48. doi: 10.1146/annurev.psych.121208.131647
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004a). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22, 1694–1703. doi: 10.1016/j.neuroimage.2004.04.015
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004b). Opposing BOLD responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport* 15, 2239–2243. doi: 10.1097/00001756-200411150-00022
- Rosenberg, S., Nelson, C., and Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *J. Pers. Soc. Psychol.* 9, 283–294. doi: 10.1037/h0026086
- Ross, L. (1977). The intuitive psychologist and his shortcomings. *Adv. Exp. Soc. Psychol.* 10, 174–221. doi: 10.1016/S0065-2601(08)60357-3
- Ross, L. R., and Nisbett, R. E. (1991). *The Person and the Situation: Perspectives of Social Psychology*. New York, NY: McGraw-Hill.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758. doi: 10.1126/science.1082976
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662. doi: 10.1016/S0896-6273(04)00014-5
- Slonim, R., and Guillen, P. (2010). Gender selection discrimination: evidence from a trust game. *J. Econ. Behav. Organ.* 76, 385–405. doi: 10.1016/j.jebo.2010.06.016
- Smith, D. V., Hayden, B. Y., Truong, T.-K., Song, A. W., Platt, M. L., and Huettel, S. A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J. Neurosci.* 30, 2490–2495. doi: 10.1523/JNEUROSCI.3319-09.2010
- Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., and Phelps, E. A. (2011). Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7710–7715. doi: 10.1073/pnas.1014345108
- Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., et al. (2012). Learning to simulate others' decisions. *Neuron* 74, 1125–1137. doi: 10.1016/j.neuron.2012.04.030
- Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131. doi: 10.1126/science.185.4157.1124
- Van't Wout, M., Kahn, R. S., Sanfey, A. G., and Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Exp. Brain Res.* 169, 564–568. doi: 10.1007/s00221-006-0346-5
- Van't Wout, M., and Sanfey, A. G. (2008). Friend or foe: the effect of implicit trustworthiness judgments in social decision-making. *Cognition* 108, 796–803. doi: 10.1016/j.cognition.2008.07.002
- Van den Bos, W., McClure, S., Harris, L. T., Fiske, S. T., and Cohen, J. D. (2007). Dissociating affective evaluation and social cognitive processes in ventral medial prefrontal cortex. *Cogn. Behav. Neurosci.* 7, 337–346. doi: 10.3758/CABN.7.4.337
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858. doi: 10.1002/hbm.20547
- Willis, J., and Todorov, A. (2006). First impressions: making up your mental state after a 100-ms exposure to a face. *Psychol. Sci.* 17, 592–598. doi: 10.1111/j.1467-9280.2006.01750.x
- Yoshida, W., Seymour, B., Friston, K. J., and Dolan, R. J. (2010). Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* 30, 10744–10751. doi: 10.1523/JNEUROSCI.5895-09.2010
- Zaki, J., Schirmer, J., and Mitchell, J. P. (2011). Social influence modulates the neural computation of value. *Psychol. Sci.* 22, 894–900. doi: 10.1177/0956797611411057

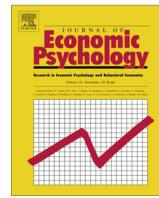
Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 October 2013; paper pending published: 21 November 2013; accepted: 10 December 2013; published online: 25 December 2013.

*Citation: Lee VK and Harris LT (2013) How social cognition can inform social decision making. *Front. Neurosci.* 7:259. doi: 10.3389/fnins.2013.00259*

*This article was submitted to Decision Neuroscience, a section of the journal *Frontiers in Neuroscience*.*

Copyright © 2013 Lee and Harris. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Affect and fairness: Dictator games under cognitive load

Jonathan F. Schulz ^{a,*}, Urs Fischbacher ^{c,d}, Christian Thöni ^b, Verena Utikal ^e

^a School of Economics, University of Nottingham, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, United Kingdom

^b University of Lausanne, Walras Pareto Center, 1015 Lausanne-Dorigny, Switzerland

^c University of Konstanz, PO Box D 131, 78457 Konstanz, Germany

^d Thurgau Institute of Economics, Hauptstrasse 90, 8280 Kreuzlingen, Switzerland

^e University of Erlangen-Nürnberg, Lange Gasse 20, 90403 Nürnberg, Germany



ARTICLE INFO

Article history:

Available online 7 September 2012

JEL classification:

C91
D03

PsycINFO classification:

2340
2360

Keywords:

Social preferences
Cognitive load
Dual-system theories
Laboratory experiment

ABSTRACT

We investigate the impact of affect and deliberation on other-regarding decisions. In our laboratory experiment subjects decide on a series of mini-Dictator games while under varying degrees of cognitive load. Cognitive load is intended to decrease deliberation and therefore enhance the influence of affect on behavior. In each game subjects have two options: they can decide between a fair and an unfair allocation. We find that subjects in a high-load condition are more generous – they more often choose the fair allocation than subjects in a low-load condition. The series of mini-Dictator games also allows us to investigate how subjects react to the games' varying levels of advantageous inequality. Low-load subjects react considerably more to the degree of advantageous inequality. Our results underscore the importance of affect for basic altruistic behavior and deliberation in adjusting decisions to a given situation.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

In 2010 David Freer risked his life to save a stranger drowning in the sea. When asked about the incident he replied: "For split a second, I thought, 'do I really want to risk stranding both of us?' Then instinct just kicked in."¹ Most theories in economics are cognitive in nature and view behavior as a deliberate act based on a thorough assessment of all possible contingencies. However, most people would agree with David Freer that instincts or affect do influence behavior – particularly in a social context.

To incorporate the role of affect, a two-system framework of the decision process has been proposed in the literature.² According to these dual process theories, two different modes of cognitive processes govern decisions: One process can be characterized as operating fast, automatically, effortlessly and often as emotionally charged. The other process operates more slowly, in a deliberate manner, and demands greater cognitive capacity. Following Loewenstein and O'Donoghue (2007) we will

* Corresponding author. Tel.: +41 (0) 21 692 28 40; fax: +41 (0) 21 692 28 45.

E-mail addresses: jonathan.schulz@unisg.ch (J.F. Schulz), urs.fischbacher@uni-konstanz.de (U. Fischbacher), christian.thoeni@unil.ch (C. Thöni), verena.utikal@wiso.uni-erlangen.de (V. Utikal).

¹ See "Father Risks his Life to Save Man in Sea" (2010).

² See for example Stanovich and West (2000), Kahneman (2003), Lieberman (2003), Strack and Deutsch (2004), Loewenstein and O'Donoghue (2007) and Evans (2008).

refer to these two modes as corresponding to affect and deliberation. This is consistent with the perspective that affect is a (fast) decision heuristic.³ A number of factors, such as situation, mood, exhaustion of willpower and cognitive load influence whether the cognitive or the affective process has a greater influence on decision-making.

In our study subjects play a series of Dictator games while they are under cognitive load. The additional memory load is intended to decrease cognitive capacity and therefore diminish deliberation.⁴ As such, decisions taken under this additional load are governed to a greater extent by the affective system. Originally introduced to study short-term memory (see Baddeley & Hitch, 1974), dual-task techniques have been successfully applied to a wide range of topics in psychological research. For example, studies show that individuals under cognitive load rely to a greater degree on stereotypes (Gilbert & Hixon, 1991), exert less self-control as measured by the choice between cake and fruit salad (Shiv & Fedorikhin, 1999) or exhibit higher discounting of future rewards (Hinson, Jameson, & Whitney, 2003).

In this paper we focus on the following question: in which direction does the affective system steer other-regarding decisions, when the deliberative system is occupied with an additional cognitive task? In other words, is fair behavior deeply rooted in humans' affective system or is it a rather effortful, cognitive process that overrides immediate selfish responses? The existing literature gives conflicting answers to this question.

One side of the debate posits that the deliberative system inhibits immediate selfish urges and guides decisions based on moral and ethical principles. The affective system – evolutionarily older and thus more related to animal behavior – is driven by immediate self-interest. The perspective that moral decisions are the result of a process of reasoning and reflection has a long history in philosophy. Like Kant's categorical imperative or the Ten Commandments, philosophy and religion offer ways of grounding values. Similarly, among evolutionary biologists, scholars have argued that civilization is only a thin veneer hiding humans' selfish nature. For example, according to Williams (1988), morality is an accidental byproduct of human evolution. This view is also reflected in Schopenhauer's (1851) quote "Man is at bottom a dreadful wild animal. We know this wild animal only in the tamed state called civilization..." or Ghiselin's (1974) "Scratch an 'altruist' and watch a 'hypocrite' bleed." In a similar vein, Moore and Loewenstein (2004) argue that self-interest is an automatic process, whereas ethical responsibilities operate via controlled processes.⁵

Contrary to this perspective, de Waal (2006) argues that human morality is more fundamental and has evolved from social instincts humans share with other animals. Support for this view comes from studies on animal behavior where basic social behavior is observed (for an overview, see Preston & de Waal, 2002). Similarly, van Winden (2007) emphasizes the importance of emotion in contrast to cognition in the individual enforcement of, as well as the compliance with, norms like fairness. According to the social intuitionist approach of Haidt (2001), moral decisions are the result of quick, automatic heuristics. His considerations are based on the observation that individuals exhibit moral reactions to hypothetical scenarios, but have difficulties in providing reasoning for their views. Empirical support for a specific heuristic – the equality heuristic – comes from Güth, Huck, and Müller (2001). Subjects were faced with mini-Ultimatum Games (UGs), where only two allocations, a fair and an unfair one, were feasible.⁶ They find that the fair allocation was chosen more often when it consisted of an equal split compared to an "almost-equal split". Their finding is in line with a focal-point interpretation: fairness concerns are only triggered when the focal equal split is feasible.

1.1. Related empirical literature

The debate on whether altruistic choice is primarily guided by deliberation or by affective reactions is far from settled. The existing evidence from neuroscience, response times and cognitive load studies is inconclusive.

Neuroscience has investigated the neural correlates of the two-system theory. Moll et al. (2006) studied charitable donations using functional magnetic resonance imaging (fMRI). They find that evolutionarily older areas of the brain associated with the affective system (mesolimbic reward system) are not only activated when receiving monetary rewards but also when giving to charity. However, brain areas associated with deliberation (areas in the prefrontal cortex) are activated when (i) individuals are opposed to the charitable cause and (ii) the decision to donate comes at a cost. This suggests that the affective system is not solely governed by material self-interest. The deliberative system on the other hand mediates the affective reaction, when it is either in conflict with more abstract moral beliefs or with self-interest. Related is the fMRI study by Sanfey, Rilling, Aronson, Nystrom, and Cohen (2003), who study the Ultimatum game. Activation in brain areas associated with the affective part of the brain (anterior insula) exhibit a positive correlation with rejection rates of unfair offers. Acceptances of unfair offers on the other hand were attributed to the cognitive part of the brain (right dorsolateral prefrontal cortex). As rejecting an unfair offer comes at a cost, their finding is further evidence that the affective system did not steer behavior towards self-interest. While these studies report correlations between behavior and brain activity, Knoch, Pasqual-Leone,

³ Thus, the focus is not on subjective feeling states associated with emotions. Following Loewenstein and O'Donoghue (2007), the defining characteristic of affect is that it carries "action tendencies". This is in contrast to expected emotions that are incorporated into deliberation.

⁴ Other studies stimulate the affective system. For example, Kirchsteiger, Rigotti, and Rustichini (2006) prime second movers by showing them either a funny or depressing movie.

⁵ See also Rachlin (2002), who views altruism as a self-control problem, and the subsequent discussion in that issue of Behavior and the Brain Science.

⁶ In a standard UG, one player, the Proposer, decides on the distribution of a sum of money. A second player, the Receiver, can either accept or reject this proposal. If accepted, the money is divided according to the proposal. If rejected, both players obtain a payoff of zero. See Güth, Schmittberger, and Schwarze (1982).

Meyer, Treyer, and Fehr (2006) investigate causal effects. They use repetitive transcranial magnetic stimulation (rTMS) to disrupt the prefrontal cortex. They find that subjects are more willing to accept unfair offers when the right prefrontal cortex is disrupted. As such their finding suggests that choices are more likely to be self-regarding when deliberation is impaired. Comparing these three studies the evidence on the role of affect on social preferences is mixed. Closest to our study is Moll et al. (2006). Like their study, our experiment is non-strategic. In contrast, however, our research has the advantage that we can draw causal inference.⁷

The existing empirical evidence on cognitive load and social preferences is likewise inconclusive. Roch, Lane, Samuelson, Allison, and Dent (2000) find that individuals under high cognitive load are more likely to request an equal split from a common resource pool. In an Ultimatum Game, Cappelletti, Güth, and Ploner (2011) do not find an effect of cognitive load. Closer to our study are the experiments by Hauge, Brekke, Johansson, Johansson-Stenman, and H. (2009), Cornelissen, Dewitte, and Warlop (2011) and Benjamin, Brown, and Shapiro (2006). All three studies focus on Dictator game giving. However, none of the studies finds a main effect of the cognitive load task that consisted of memorizing a seven digit-number. Cornelissen et al. (2011) find a treatment difference for a subset of individuals – those that were classified as pro-socials in a different task give a higher amount in the high-load condition.⁸

In this paper we focus on Dictator games (DGs). Compared to previous research on cognitive load and DG, our experimental study comprises two main innovations: First, we apply a different cognitive load task. Hauge et al. (2009) suggest that their cognitive load task (seven-digit number) might have been insufficient to find treatment effects. We impose cognitive load by an *n*-back task (Gevins & Cutillo, 1993). Subjects hear a series of letters and have to press a button whenever they hear a character that resounded two letters before. This task is likely to impose a higher cognitive load than simple tasks like memorizing seven digit numbers. In addition to solely memorizing, *n*-back tasks require monitoring, updating and manipulation of information. *n*-Back tasks have been used in *functional magnetic resonance imaging* (fMRI) and positron emission tomography (PET) studies to investigate the role of working memory. They have consistently shown increasing activity of the frontal cortex (for overviews see Fletcher & Henson, 2001; Owen, McMillan, Laird, & Bullmore, 2005). Thus, our memory load task has been shown to specifically activate those areas in the brain that are associated with deliberation. Second, our experiment comprises a series of mini-DGs. This allows us (i) to investigate how subjects react to varying incentives posed by different mini-DGs and (ii) test the equality heuristic. The implementation of a series of games also has a methodological purpose. Other studies informed subjects on their whole choice set before they were under cognitive load. In principle, subjects in these studies could decide prior to being under load. In our experiment we informed subjects only about the general structure of the decision situation. The payoff structure of the particular game was revealed under cognitive load. Thus, subjects could not choose their strategy prior to being under cognitive load.

2. Experimental design and procedures

Our experiment consists of two parallel tasks. While subjects are engaged in a cognitive load task they simultaneously decide on a social task. Our treatment variation is the difficulty of the cognitive load task. Subjects are randomized to either a high- or a low-load condition.

2.1. Social decision task

The social decision task consists of a series of 20 mini-Dictator games (mini-DGs). In each mini-DG the Dictator decides on the distribution of money between himself and an anonymous other. The choice set is restricted to two allocations. One allocation always exhibits a greater inequality (unfair allocation) than the other (fair allocation). Table 1 lists the 20 mini-DGs. For example, in Game No. 1 subjects can decide between the allocation 50/50 and 60/40.

Apart from the overall effect of cognitive load, this series of mini-DGs allows us to investigate how individuals in the two treatment conditions react to varying degrees of inequality in the various games. For example, the unfair allocation in the mini-DG with the allocations 50/50 and 60/40 (Game No. 1) leads to less inequality (and lower payoff to the Dictator) than in the game with 50/50 and 80/20 (Game No. 9): in the former game the Receiver gets twenty points less than the Dictator, whereas in the latter it is 60 points. We hypothesize that subjects under low load are more responsive to the different incentives posed as they have more cognitive resources to evaluate each single game.

For every mini-DG with an equal split we included an additional one with an “almost-equal split” slightly favoring the Dictator. This allows us to test the hypothesis that an equal split constitutes a focal point as suggested by Güth et al. (2001), Roch et al. (2000), or Messick and Schell (1992). Thus, if the equal split constitutes a decision heuristic, we would

⁷ See also Rubinstein (2007) and Piovesan and Wengström (2009) on response times and social preferences. The latter find longer response times for pro-social choices in a version of a Dictator game. As far as longer response times reflect more cognitive activity, their results suggest that it is deliberation overriding immediate selfish responses. In a strategic situation Rubinstein (2007) finds the opposite. Egoistic decisions of Proposers in the Ultimatum game exhibit longer reaction times. An early study on time pressure and helping behavior is Darley and Batson (1973).

⁸ See also the studies by van den Bos, Peters, Bobocel, and Ybema (2006) and Skitka, Mullen, Griffin, Hutchinson, and Chamberlin (2002) which investigate subjects' evaluation of hypothetical scenarios under varying load conditions. Van den Bos et al. (2006) report that high-load subjects express a higher level of satisfaction with advantageous inequality, while Skitka et al. (2002) find for a subset of subjects (liberalists) that they are less willing to help someone in need when they are cognitively busy. Related is also the study by Barnes, Schaubroeck, Huth, and Ghumman (2011). They show that low levels of sleep (which is negatively related to self-control resources) is positively related to unethical behavior like cheating.

Table 1
The 20 mini-Dictator games.

Game No.	Core-games allocation, rounded		Equal Spilt	Pie size	Fair Allocation		Unfair Allocation	
					Dictator	Receiver	Dictator	Receiver
1			Equal	100	50	50	60	40
2				94	47	47	56	38
3		50–60	Not eq.	100	51	49	60	40
4				94	48	46	56	38
5			Equal	100	50	50	70	30
6				94	47	47	66	28
7		50–70	Not eq.	100	51	49	70	30
8				94	48	46	66	28
9			Equal	100	50	50	80	20
10				94	47	47	75	19
11		50–80	Not eq.	100	51	49	80	20
12				94	48	46	75	19
13			Equal	100	50	50	90	10
14				94	47	47	85	9
15		50–90	Not eq.	100	51	49	90	10
16				94	48	46	85	9
17		80–90	Not eq.	100	80	20	90	10
18				94	75	19	85	9
19		60–100	Not eq.	100	60	40	100	0
20				94	56	38	94	0

Note: The social decision task consists of a series of 20 binary mini-DGs. Column 2: the game's varying degree of inequality; the first number refers to the (rounded) Dictator's percentage share in the fair, the second in the unfair allocation. Column 3: for each game with an equal split we included one with an almost equal split. Column 4: for each game with a pie-size of 100 we included an otherwise identical one with a pie-size of 94. Column 5: amount of points to the Dictator and Receiver in the Fair-Allocation. Column 6: amount of points to the Dictator and Receiver in the unfair Allocation.

expect to see a higher percentage of individuals choosing the equal split compared to the almost-equal split in otherwise identical mini-DGs. This effect should be exaggerated under high cognitive load, as the decisions are less influenced by deliberation.

Each mini-DG has a counterpart exhibiting a slightly different pie-size. In particular, 10 games exhibit a pie-size of 100 and 10 games a pie-size of 94. The relative shares in the respective games are identical up to rounding differences. We did this to investigate possible heuristics and as robustness check. A pie-size of 100 may be more easily accessible than a pie-size of 94 as the percentage shares and levels coincide in the former case. For example, general linguistic usage denotes an equal split as a fifty-fifty option. In case of a pie-size of 100 the equal split corresponds to 50 points each. Therefore, it might constitute a stronger focal point than the equal split of 47 points each. A similar argument can be made for the other allocations.

To conclude, our experimental design consists of four “core-games”, each coming in four flavors differing along two dimensions: (i) the pie-size and (ii) whether the fair allocation constitute an equal split or an almost equal split (Games No. 1–16 in Table 1). We included two more core-games without an equal or almost equal split (Games No. 17–20). The only variation within these core-games is the pie-size. On the one hand, we were interested in how behavior is affected when the fair allocation exhibits a greater degree of inequality. On the other hand, we wanted to introduce more variation in our games so that the systematic design of our games does not become too obvious for the subjects.

2.2. Cognitive load task

Our cognitive load task consists of an *n*-back task. In our *n*-back task subjects hear a new letter over headphones every three seconds. In the high-load condition, subjects are incentivized to press a key every time they hear a letter that resounded two letters before (2-back condition). In the low-load condition (0-back) subjects are incentivized to indicate every time they hear the letter “L”. Altogether the sequence consists of 10 different letters (D, F, K, L, N, P, Q, R, S, T) and 25% are targets, that is, letters to be indicated. The letters are recorded in one female and one male voice and sound in randomized order. The sequence is constructed such that in both load conditions the targets occur at the same time. For every correct indication of a target subjects receive 0.5 points. If a subject indicates incorrectly, 0.25 points are deducted. Parallel to the cognitive load task they complete the social decision task. Jaeggi et al. (2003) have shown that subjects are capable of completing two parallel tasks – in their study two 2-back tasks – and perform well above chance.

2.3. Procedures

We conducted five sessions with 136 participants in June and July 2010 at the LakeLab of the University of Konstanz. Participants were students of the University of Konstanz and were recruited using the online recruiting system ORSEE (Greiner, 2004). None of the subjects participated in more than one session. Each subject sat at a randomly assigned PC terminal and

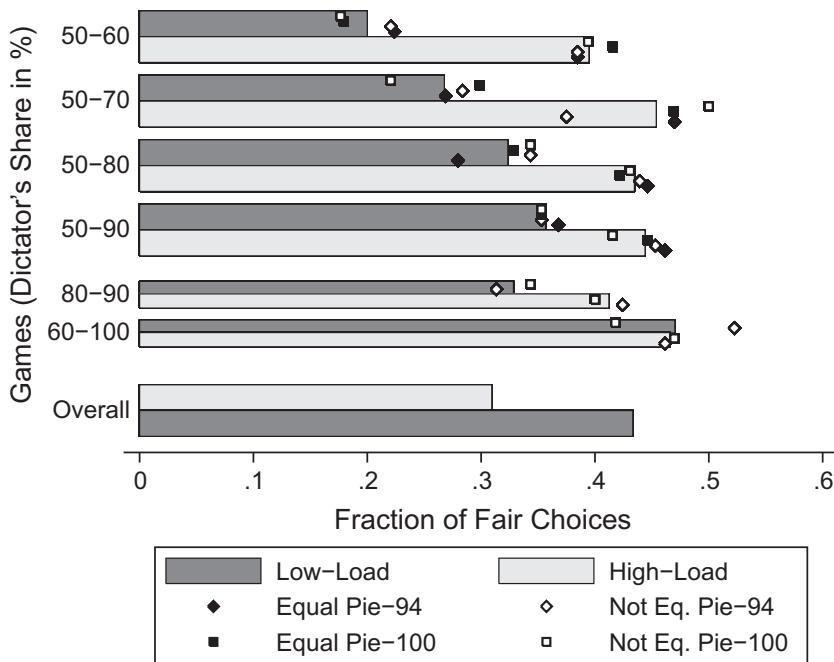


Fig. 1. Fraction of fair choices by load and game. The first number on the y-axis refers to the (rounded) Dictator's share in the fair-allocation, the second number to the share in the unfair allocation. Thus, Games with the same (rounded) fraction of the unfair allocations are pooled (those that differ only in pie-size and whether or not fair allocation is equal or almost equal split). Please see Table 2 for the regression analysis, which allows controlling for the period the decision was taken.

was given a copy of instructions (see the online supplementary material for the instructions). The experiment was programmed and conducted with the software *z-Tree* (Fischbacher, 2007). A set of control questions was provided to ensure the understanding of the game. The experiment did not start until all subjects had answered all questions correctly. In order to ensure the understanding of the *n*-back task participants took part in an unpaid practice round for 90 s.

The order of the 20 mini-DGs was randomized. Subjects had 20 s to decide in a mini-DG followed by a 7 s break before a new game started. Parallel to it they took part in the cognitive load task. All subjects took decisions as a Dictator. Only at the end of the experiment was the actual role of a participant (either Dictator or Receiver) randomly determined. Further, only one randomly determined game was paid out. Thus, 50% of (randomly determined) participants were paid according to their decision in one (randomly determined) mini-DG as the Dictator. The other 50% were Receivers of the corresponding games. One point of the randomly chosen game translated into 0.22 €. Average income amounted to about 17 €: 10.5 € for the social decision task, 4.5 € for the *n*-back task (2 € for show-up). The experiment itself lasted about 70 minutes.

3. Results

Our cognitive load treatment is only effective if subjects actually exert effort in our *n*-back task. We find that this is indeed the case: the performance (percentage of non-missed targets and no wrong indication) was 99.6% in the 0-back and 96.0% in the 2-back condition. This suggests that the 2-back task is more demanding, but people still complete it well above chance. Taken together these results indicate that subjects were successfully put under cognitive load.

Focusing on the treatment differences we find that subjects in the high-load condition are more generous on average. They choose the fair allocation 43.3% of the time compared to 30.9% in the low-load condition. This treatment difference suggests that once the affective system is mediated to a lesser extent by the deliberative system, choices are more generous.⁹

Looking at core games reveals interesting heterogeneity in the treatment effect. The games vary in their extent of inequality (and hence payoff) of the two allocations. Fig. 1 displays the fraction of fair choices for every core game by cognitive load (bars). Dots show the results of the four individual games. In almost every game the fraction of fair choices is larger in the high-load condition. Treatment differences are more pronounced in games exhibiting only a small level of inequality.

As it is apparent from Fig. 1, individuals in the low-load condition react more strongly to the incentives posed by the different games. The larger the inequality of the unfair allocation, the more likely are low-load subjects to choose the fair

⁹ If a subject does not enter a decision within the 20 s then the observation is missing for the particular social decision task. Altogether, 2.6% of the decisions are missing. In the high-load condition this was the case 4.3% of the time, largely reflecting that one subject in the high load condition did not make any choice at all.

Table 2

Probit regression of fair choices on load.

	(1) Fair Option	(2) Fair Option	(3) Fair Option	(4) Fair Option
Highload	0.139* (0.074)	0.507*** (0.106)	0.481*** (0.107)	0.498*** (0.107)
Fraction Dictator Unfair Allocation		0.659*** (0.113)	0.656*** (0.112)	0.659*** (0.113)
(Fraction adv) × highload		-0.516*** (0.139)	-0.509*** (0.138)	-0.516*** (0.139)
Fraction Dictator Fair Allocation		-0.15 (0.108)	-0.163 (0.111)	-0.149 (0.108)
(Fraction fair.) × highload		0.018 (0.131)	0.054 (0.134)	0.018 (0.131)
Equal Option			-0.007 (0.02)	
(Equal Option) × highload			0.018 (0.024)	
Pie-Size 100				-0.016 (0.017)
(Pie-Size 100) × highload				0.022 (0.021)
Period	0.002 (0.002)	0.003 (0.002)	0.003 (0.002)	0.003 (0.002)
Period × highload	-0.001 (0.003)	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)
N	2650	2650	2650	2650
Pseudo R ²	0.013	0.024	0.024	0.023

Note: Marginal effects of probit estimation with robust standard errors clustered on subjects in parenthesis. The dependant variable is a dummy indicating a fair choice. × denotes interaction terms. 'Fraction Dictator Unfair Allocation' denotes the number of points to the Dictator in the unfair allocation. The dummy 'Equal Option' denotes whether the fair allocation is an equal split. 'Pie-Size 100' indicates whether the pie-size is 100. 'Fraction Dictator Fair Allocation' denotes the number of points to the dictator in the fair allocation.

* Significant at $p < 0.1$.

**Significant at $p < 0.05$.

*** Significant at $p < 0.01$.

allocation. For example, only 20% of low-load subjects decide for the fair allocation when the unfair allocation leaves the Receiver 40% of the pie. However, 35.7% choose the fair allocation when the choice is between the equal split and leaving 10% for the Receiver.

The probit regression in **Table 2** corroborates these findings. Without controlling for the inequality of the different allocations, high-load subjects choose the fair allocation weakly significantly more often (column 1).¹⁰ Conditioning on the degree of inequality reveals that the two load conditions are highly significantly different: low-load subjects react significantly stronger to the incentives posed by the inequality of the unfair allocation (column 2). An increase in the Dictator's share in the unfair allocation by 10 percentage points leads to an increase in the probability of a fair choice by 6.6 percentage points. Even though high-load subjects react significantly less, an F-test of joint significance reveals that they also weakly significantly react to the inequality of the unfair allocation ($p = 0.087$). As a result, cognitive load is more effective in situations where the unfair allocation exhibits a relative small degree of inequality. Our findings therefore suggest that subjects under low-load on average take the situation more fully into account. If the unfair allocation leads to only a small degree of inequality, they behave in a self-interested way. However, in more extreme instances, in particular when the unfair allocation leaves nothing for the other individual, the treatment difference vanishes.

3.1. Equality heuristic and pie-size

In our experiment we do not find evidence for an equality heuristic. There appear to be no systematic differences based on whether the fair allocation is the equal split or the almost-equal split (see **Table 2** column 3). This is also the case when focusing on high-load subjects only. If the equality heuristic exists, we would expect a more pronounced effect when subjects are cognitively busy. However, as **Fig. 1** and **Table 2** (column 3) reveal, high-load subjects are also not more likely to choose an equal split over an almost-equal split. This suggests that the equal split does not constitute a focal point in our experiment. Compared to the study by Güth et al. (2001), our social decision task is non-strategic. In a strategic setting such as the Ultimatum Game the equal split might be an attractive choice due to the (beliefs about) behavior of second movers. Similarly, we do not find an effect of the different pie-sizes (**Table 2**, column 4). Thus, whether the actual points coincide with the percentage distribution or not seems to have no effect on the outcome. In all regressions we also control for the time at which a particular mini-DG was played during the experiment (Period, and interaction with HighLoad). As **Table 2** reveals we do not find a time trend in our data: decisions do not systematically vary with the variable 'Period'.

3.2. Individual decisions

Individuals under high-load are more generous and react less to the incentives posed by the different games. Does this simply reflect a higher degree of randomness in subjects' decisions? That is, if individuals in the high-load condition are more likely to make random decisions, the mean will be closer to the expected random outcome of 0.5.

To test whether high-load individuals exhibit a higher degree of randomness we consider individual decisions. A benchmark for consistency has to specify the impact of own and the other's payoff on utility in a coherent way. In our non-strategic setting outcome based models of inequality aversion offer a point of departure. In the Fehr and Schmidt (1999) model utility

¹⁰ The treatment effect becomes significant at $p < .05$ if we restrict the sample to the four core games with an equal or almost equal split.

Table 3

Number of subjects by (consistent) strategies and treatment condition (in the 16 games with equal or almost equal split).

	Number of subjects low-load	Number of subjects high-load
Always fair	4 (5.9%)	12 (17.9%)
Never fair	33 (48.5%)	28 (41.8%)
Become fair as unfair allocation exhibits higher inequality	9 (13.2%)	7 (10.4%)
Become Egoistic as unfair allocation exhibits higher inequality	1 (1.5%)	1 (1.5%)
Rest	21 (30.9%)	19 (28.4%)
Mean pseudo- R^2 (rest)	0.12	0.10
Mean pseudo- R^2 (all subjects)	0.73	0.74

Note: Mean pseudo- R^2 (rest) denotes the mean of pseudo- R^2 s obtained from individual Probit estimations on the subject level of those that do not have a consistent strategy. To calculate Mean pseudo- R^2 (all subjects) the R^2 of subjects exhibiting a consistent strategy were set to 1.

depends linearly on advantageous inequality. Therefore, the model predicts that subjects either always choose the fair allocation or always choose the unfair allocation. Fehr-Schmidt are aware that their assumption of linearity is not fully realistic – especially in the DG. They acknowledge that a non-negligible fraction of people exhibit nonlinear inequality aversion in the domain of advantageous inequality.

We therefore turn to Bolton and Ockenfels (2000). In their model utility is nonlinear in inequality aversion. Utility is convex in inequality and as a result the model exhibits an increasing marginal sensitivity towards inequality. Thus, in the standard DG it does not restrict optimal choices to the equal split or the pure selfish allocation, but supports all allocations in between. What are the implications for the mini-DGs? It is straightforward that individuals, who either always or never choose the unfair allocation, reveal consistent behavior. That is, for “always-fair” individuals in each binary game the monetary gain of the unfair allocation is lower than the implied (psychological) loss due to inequality.

For individuals who switch between fair and unfair allocations, Bolton-Ockenfels gives straightforward predictions if we solely focus on the 16 allocation decisions with an equal (and almost-equal) split.¹¹ Restricting the analysis to these 16 games allows us to focus solely on the varying degree of inequality of the unfair allocation: at a certain threshold – as the inequality of the unfair allocation increases – individuals previously choosing the unfair allocation switch to the fair allocation. Up to the threshold the monetary gain dominates the (psychological) losses from inequality. Past the threshold, inequality aversion dominates the monetary gains (in an unrestricted choice set their DG choice would lie somewhere close to this threshold).

The increasing marginal sensitivity towards inequality of Bolton–Ockenfels implies this pattern where individuals switch from the unfair allocation to the fair allocation as inequality gets larger. However, it also seems plausible that individuals exhibit decreasing marginal sensitivity, that is, subjects consistently switch from the fair to the unfair allocation as inequality gets larger. We therefore included this possibility in our analysis.

According to this measure 70.6% of the individuals in the high-load condition and 67.1% in the low-load condition behave in a manner consistent with Bolton–Ockenfels utility functions. Thus, there are almost no differences in our consistency measure of the two treatment groups. The largest fraction consists of individuals who never choose the fair allocation in the 16 games with an equal or almost-equal split. As Table 3 reveals “neverfair” makes up a larger fraction (48.5%) in the low-load condition compared to the high-load condition (41.8%). In the high-load condition in contrast a considerably larger number of subjects always choose the fair allocation (17.9%) compared to (5.9%) in the low-load condition.¹² To get a measure for consistency of the remaining subjects we estimated individual probit regressions. This was done by regressing the 20 choices of one individual on the extent of inequality (that is, the Dictator's share) of the unfair allocation. The resulting individual pseudo- R^2 gives an indicator of consistency. There are only minor differences in the two means of the pseudo- R^2 s (see Table 3) of the subjects, which do not exhibit a consistent pattern. A Wilcoxon Signed-Rank test reveals that they are not significantly different ($p = 0.85$). This suggests that individuals in the high-load condition exhibit behavior that is as consistent as in the low-load condition. Therefore, the finding that high-load subjects are more generous is unlikely to reflect a higher degree of randomness in subjects' choices. Our result rather shows that a higher fraction of high-load subjects always choose the fair allocation, whereas low-load subjects are more likely to never choose the fair allocation. Cognitive load in Mini-DGs does not cause random behavior but makes participants more generous.

4. Conclusion

In which direction does the affective system steer other-regarding decisions, when the deliberative system is occupied with an additional cognitive task? Utilizing a dual-task technique we find that individuals' choices are more generous when taken under high cognitive load.

¹¹ For our consistency measure we only focus on the degree of inequality of the unfair allocation. That is, we neglect differences in the fair allocation stemming from equal split and almost equal split. Additionally, we focus only on the relative distribution of the advantageous allocation, that is, we neglect the minor differences stemming from the two different pie-sizes. Incorporating these differences does not lead to any qualitative changes of the results.

¹² As we randomized the appearance of the fair and advantageous allocations (up or down), always choosing the fair or advantageous allocation does not constitute an easy heuristic like “always choose the upper allocation”.

This finding underscores the importance of the affective system in basic altruistic choices. Our evidence suggests that the affective system steers behavior towards altruistic choice.¹³ Thus, our study supports the notion that basic social preferences are fundamental: the affective system, associated with heuristics and evolutionarily older parts of the brain, mediates decisions towards altruistic choices. This suggests that basic morality is an (older) product of evolution and not just a “thin layer of civilization covering the wild animal within”. While evolutionary theory posits a selfish gene (Dawkins, 1976) this does not have to lead to selfish behavior. Kin-selection (Hamilton, 1964), reciprocity (Trivers, 1971), indirect reciprocity (Alexander, 1987; Nowak & Sigmund, 1998), costly signaling (Gintis, Smith, & Bowles, 2001) and gene-culture coevolution (Gintis, 2003) can explain cooperative behavior. Affective reactions might be an important proximal mechanism for supporting cooperative behavior in these instances.

The studies by Kogut and Ritov (2005) and Small, Loewenstein, and Slovic (2007) are related. They find that subjects exhibit a higher willingness to donate to identifiable victims. They attribute this to the role of emotions. In fact, Small et al. (2007) show that inducing people to deliberate about the discrepancy in giving towards identifiable and statistical victims results in an overall reduction in donations. However, even though related, we do not view the affective system (or affect) as a synonym for emotions. Our results relate more to the interpretation that basic altruism is a fast decision heuristic. Nevertheless this does not exclude (immediate) emotions as the driving force behind these decision heuristics.¹⁴ In our social decision task subjects are placed in an environment where emotions like empathy are a likely factor explaining our results.

Our experiment highlights the importance of the deliberative system. In the low-load condition individuals react more strongly to incentives posed by differences in the inequality of the various games. This lends credence to the interpretation that inequity aversion seems to require more cognitive resources than simple generous behavior. It involves a more thorough welfare assessment and comparison. Thus, while the deliberative system adjusts behavior in a self-serving manner, it also moderates the immediate affective reaction in a way that is more tailored to the situation at hand. For example, in the case that the unfair allocation does not leave any points for the other person, the low-load subjects are just as likely to choose the fair allocation.

We find no indication for an equality heuristic in our experiment. Individuals are just as likely to choose an equal split or an almost-equal split. Thus, in our study the affective system more generally steers towards altruistic behavior and this is not reflected by the focal point of an exact equal split. In our non-strategic experiment subjects were confronted with rather straightforward social dilemmas. How our results extend to more complex moral settings (e.g. third party inequalities as in Johansson & Svendsäter, 2009) or situations that trigger emotions like anger or envy might be a worthwhile area for future research.

Acknowledgments

Jonathan F. Schulz and Christian Thöni gratefully acknowledge financial support from the Richard Büchner Stiftung. Urs Fischbacher gratefully acknowledges financial support from the German Research Foundation (DFG) through research unit FOR 1882 “Psychoeconomics”. Verena Utikal gratefully acknowledges financial support from the Center of Psychoeconomics at the University of Konstanz. We thank Marco Piovesan, Jean-Robert Tyran, two anonymous referees and participants at the Thurgau Experimental Economics Meeting 2011 and the IAREP/SABE//ICABEEP conference 2011 for valuable comments. We thank Jean-Robert Tyran for the generous support of our research at the Center for Experimental Economics at the University of Copenhagen.

Appendix A. Instructions to participants (translated from German) and figures in line with text

General Instructions

Today you are participating in an economics experiment. By carefully reading the following instructions, you can – depending on your decisions – earn money in addition to the show-up fee of 2 euros. It is, therefore, of importance that you accurately read these instructions. During the whole experiment communication with other participants is not allowed. Therefore, we ask you to not speak with each other. If you do not understand something, please consult the instructions again. If you still have questions, please raise your hand. We will come to your place and answer your question individually.

During the experiment we do not speak of euros, but points. The points you earn during the experiment will be converted at the following rate:

$$1 \text{ point} = € 0.22$$

The show-up fee of 2 euros and the total number of points you earned will be converted into euros and paid out to you in cash at the end of today's experiment.

On the following pages we explain the course of the experiment in detail. First, we will familiarize you with the basic decision situation. When you are finished reading the instructions, you will find control questions on your screen. They

¹³ Since our load task manipulates solely cognition, our study cannot directly prove that decisions are more affective. For this claim we have to rely on the existing cognitive load literature, which shows that decisions under cognitive load are governed to a greater extent by affect or fast decision heuristics.

¹⁴ Loewenstein, Weber, Hsee, and Welch (2001) distinguish between immediate and expected emotions. While immediate emotions relate to affect, expected emotions are anticipated and therefore incorporated into deliberation. Similarly, physiologically Bechara and Damasio (2005) stress the importance of the prefrontal cortex (the deliberative system) in the ability to express emotions and experience feelings.

are intended to help you understand the setting. The experiment only begins when every participant is familiarized with the course of the experiment.

The experiment involves two types of participants: participant A and participant B. Participant A takes several decisions. Participant B makes no decision. Each participant takes on the role of a participant A and the role of a participant B. At the end of the experiment it will be randomly determined whether you will be paid out as a participant A or a participant B. At no point in time will you be informed about the identity of another participant. Likewise, the other participants will not be informed about your identity. Thus, all payments will be made anonymously. That is, the other participants do not learn how much you earned in the experiment.

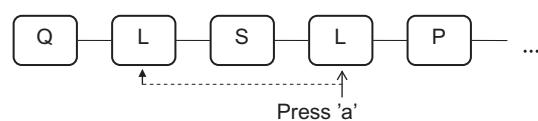
The experiment

The experiment consists of two different tasks. The first task is a listening task. Here you can earn points by responding correctly to letters you hear over the headphones. The second task consists of a sequence of 20 decision situations. In each decision situation you decide on the distribution of an amount of money between you and participant B.

Listening task

In the listening task you hear letters over your headphones. Every three seconds you hear the next letter. Your task is to press the key 'a' whenever a letter resounds that sounded 2 letters before.

Assume, for example, you hear the following sequence of letters: Q, L, S, L, P, Q, P... When you hear one of the underlined letters, you should indicate so by pressing the key 'a'.



Every time you correctly identify a letter that sounded two letters before you earn 0.5 points. To press the 'a' key you have time until the next letter sounds (3 s). If you press 'a', even though the letter did not sound two letters before, 0.25 points are deducted.

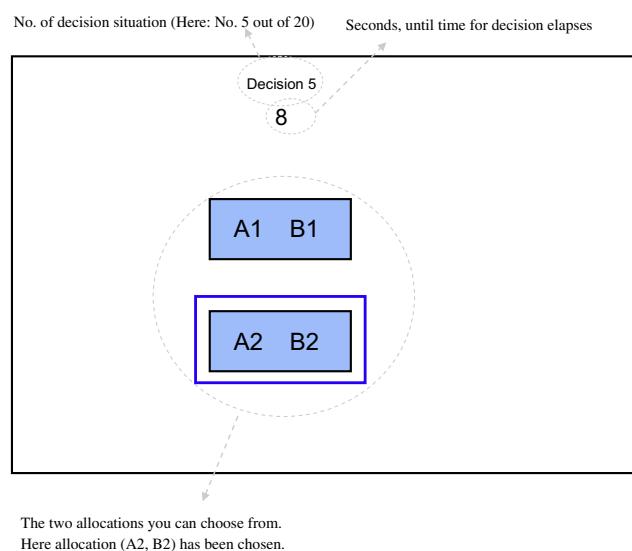
Before the experiment starts, there will be a test trial so you can familiarize yourself with the task. The test trial lasts 90 s. In this test trial there are no points to be earned. If you have any question after the test trial please do not hesitate to direct them to us by raising your hand. Please note that the use of any writing utensils during the experiment is not allowed.

Decision situations

There are 20 decision tasks. Your task is to decide on one out of two possible allocations. By deciding on one allocation, you decide how an amount is divided between you and another participant.

Display on the screen

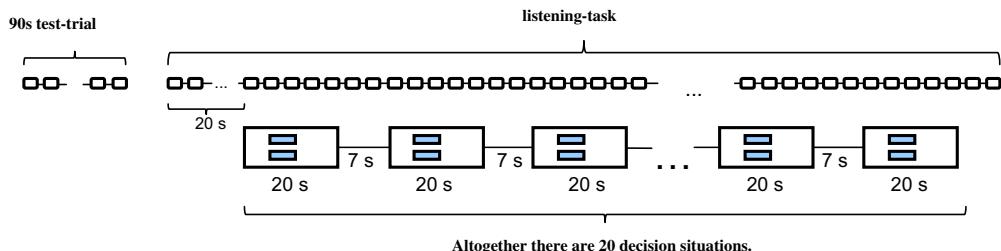
Instead of A1, B1, A2 and B2 you will find numbers which correspond to the payments to A and B. You make your choice by clicking with the left mouse button on one of the two (light blue) allocations. You have the choice between (A1, B1) and (A2, B2). If, for example, you choose (A1, B1), you propose an assignment of points in such a way that you get A1 points and participant B gets B1 points. (The left number will always refer to the number of points for you and the right number refers to the number of points to participant B). The allocation you have chosen will be highlighted with a blue rectangle. Participant B does not make a decision.



You have 20 s to make your choice. Within the 20 s you still have the opportunity to change your mind. After 20 s the highlighted allocation is taken as your choice. If you fail to make a choice in the given time, 1 point of your earnings will be deducted. How many seconds are left for your decision is shown on the screen. The number of the current decision situation is also displayed.

Sequence of the experiment

After you have read the instructions there will be a test trial of the listening task (90 s). During the test trial you cannot earn any points. There will also be control questions with regard to the decision situations. Please do not hesitate to direct any question to us.



The study starts with the listening task. The listening-task will continue throughout the study. For every correct hit you earn 0.5 points, while for every wrong hit 0.25 points are deducted. You learn the number of points you earned only at the end of the study.

Shortly after the listening task has started, the sequence of 20 decision situations begins parallel to it. In each decision situation you have 20 s to decide. Before the next decision situation starts, there is a seven second break. In each decision problem you are randomly rematched with a participant B.

Payment

At the end of the study you will be informed on the number of points you get from the listening task and the decision situations (as participant A and participant B). Your payment consists of your show-up fee (2 €), plus the number of points from the decision situation and the listening task. At the end of the experiment it will be randomly determined which decision situation will be paid out. Further, it will be randomly determined whether you will be paid out in the role as participant A or participant B. The points you earned will be converted into euros.

Control questions

Before we begin with the experiment please answer a few questions on the computer screen. These control questions do not influence your payments at the end of the experiment. First, there will be questions regarding the decision situations. When all participants have solved these questions, there will be the subsequent trial of the listening task (90 s).

References

- Alexander, R. D. (1987). *The biology of moral systems*. New York: Aldine De Gruyter.
- Baddeley, A., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). New York: Academic Press.
- Barnes, C. M., Schaubroeck, J., Huth, M., & Ghuman, S. (2011). Lack of sleep and unethical conduct. *Organizational Behavior and Human Decision Processes*, 115, 169–180.
- Bechara, A., & Damasio, A. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, 52, 336–372.
- Benjamin, D. J., Brown, S. A. & Shapiro, J. M., (2006 May 5). Who is “behavioral”? Cognitive ability and anomalous preferences. Retrieved from <<http://ssrn.com/abstract=675264>>.
- Bolton, G. E., & Ockenfels, A. (2000). ERC – A theory of equity, reciprocity and competition. *American Economic Review*, 90, 166–193.
- Cappelletti, D., Güth, W., & Ploner, M. (2011). Being of two minds: An ultimatum experiment investigating affective processes. *Journal of Economic Psychology*, 32, 940–950.
- Cornelissen, G., Dewitte, S., & Warlop, L. (2011). Are social value orientations expressed automatically? Decision making in the dictator game. *Personality and Social Psychology Bulletin*, 37, 1080–1090.
- Darley, J. M., & Batson, C. D. (1973). “From Jerusalem to Jericho”: A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27, 100–108.
- Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
- de Waal, F. (2006). *Primates and philosophers. How morality evolved*. Princeton: Princeton University Press.
- Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–278.
- Father Risks his Life to Save Man in Sea (2010 October 1). *Exmouth Journal* 24. Retrieved from <http://www.exmouthjournal.co.uk/news/father_risks_his_life_to_save_man_in_sea_1_659849> Retrieved 29.07.11.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114, 817–868.
- Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, 171–178.
- Fletcher, P., & Henson, R. (2001). Frontal lobes and human memory: Insights from functional neuroimaging. *Brain*, 124, 849–881.
- Gevins, A. S., & Cutillo, B. C. (1993). Neuroelectric evidence for distributed processing in human working memory. *Electroencephalography and Clinical Neurophysiology*, 87, 128–143.
- Ghiselin, M. T. (1974). *The economy of nature and the evolution of sex*. Berkeley, CA: University of California Press.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*, 60, 509–517.
- Gintis, H. (2003). The hitchhiker's guide to altruism: Gene-culture coevolution, and the internalization of norms. *Journal of Theoretical Biology*, 220, 407–418.

- Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology*, 213, 103–119.
- Greiner, B. (2004 November 13). *The online recruitment system Orsee 2.0 – A guide for the organization of experiments in economics*. University of Cologne, Working Paper Series in Economics, No. 10. Retrieved from <<http://sourceforge.net/projects/orsee/files>>.
- Güth, W., Huck, S., & Müller, W. (2001). The relevance of equal splits in ultimatum games. *Games and Economic Behavior*, 37, 161–169.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3, 367–388.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814–834.
- Hamilton, W. D. (1964). Genetical evolution of social behavior I, II. *Journal of Theoretical Biology*, 7, 1–52.
- Hauge, K. E., Brekke, K. L., Johansson, L.-O., Johansson-Stenman, & Svärd, H. (2009 July). Are social preferences skin deep? Dictators under cognitive load. University of Gothenburg, Working Papers in Economics, No. 371. Retrieved from <http://gupea.ub.gu.se/bitstream/2077/20768/1/gupea_2077_20768_1.pdf>.
- Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 298–306.
- Jaeggi, S. M., Seewer, R., Nirkko, A. C., Eckstein, D., Schroth, G., & Groner, R. (2003). Does excessive memory load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: Functional magnetic resonance imaging study. *NeuroImage*, 19, 210–225.
- Johansson, L.-O., & Svärd, H. (2009). Piece of cake? Allocating rewards to third parties when fairness is costly. *Organizational Behavior and Human Decision Processes*, 109, 107–111.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93, 1449–1475.
- Kirchsteiger, G., Rigotti, L., & Rustichini, A. (2006). Your morals might be your moods. *Journal of Economic Behavior and Organization*, 59, 155–172.
- Knoch, D., Pasqual-Leone, A., Meyer, K., Treyer, V., & Fehr, E. (2006). Diminishing reciprocal fairness by disrupting the right prefrontal cortex (vol. 314). New York, NY: Science, pp. 829–832.
- Kogut, T., & Ritov, I. (2005). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes*, 97, 106–116.
- Lieberman, M. (2003). Reflexive and reflective judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. D. Williams, & W. Von Hippel (Eds.), *Social judgments: Explicit and implicit processes* (pp. 44–67). New York: Cambridge University Press.
- Loewenstein, G., & O'Donoghue, T. (2007 June). *The Heat of the moment: Modeling interactions between affect and deliberation*. Retrieved from <<http://wwwarts.cornell.edu/econ/ed01/heat.pdf>>.
- Loewenstein, G., Weber, E., Hsee, C., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267–286.
- Messick, D. M., & Schell, T. (1992). Evidence for an equality heuristic in social decision making. *Acta Psychologica*, 80, 311–323.
- Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., & Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 15623–15628.
- Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research*, 17, 189–202.
- Nowak, M., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393, 573–577.
- Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25, 46–59.
- Piovesan, M., & Wengström, E. (2009). Fast or fair? A study of response times. *Economics Letters*, 105, 193–196.
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *The Behavioral and Brain Sciences*, 25, 1–20. discussion 20–71.
- Rachlin, H. (2002). Altruism and selfishness. *The Behavioral and Brain Sciences*, 25, 239–250. discussion 251–296.
- Roch, S., Lane, J. A. S., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organizational Behavior and Human Decision Processes*, 83, 185–212.
- Rubinstein, A. (2007). Instinctive and cognitive reasoning: A study of response times. *The Economic Journal*, 117, 1243–1259.
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). *The neural basis of economic decision-making in the ultimatum game* (vol. 300). New York, NY: Science, pp. 1755–1758.
- Schopenhauer, A. (1851). *Essays and aphorisms (Parerga and Paralipomena)*. New York: Penguin Books.
- Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26, 278–292.
- Skitka, L. J., Mullen, E., Griffin, T., Hutchinson, S., & Chamberlin, B. (2002). Dispositions, scripts, or motivated correction? Understanding ideological differences in explanations for social problems. *Journal of Personality and Social Psychology*, 83, 470–487.
- Small, D. A., Loewenstein, G., & Slovic, P. (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and the Human Decision Processes*, 102, 143–153.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *The Behavioral and Brain Sciences*, 23, 645–665.
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, 8, 220–247.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46, 35–57.
- van den Bos, K., Peters, S. L., Bobocel, D. R., & Ybema, J. F. (2006). On preferences and doing the right thing: Satisfaction with advantageous inequity when cognitive processing is limited. *Journal of Experimental Social Psychology*, 42, 273–289.
- van Winden, F. (2007). Affect and fairness in economics. *Social Justice Research*, 20, 35–52.
- Williams, G. C. (1988). Reply to comments on "Huxley's evolution and ethics in sociobiological perspective". *Zygon*, 23, 437–438.

Neural foundations to moral reasoning and antisocial behavior

Adrian Raine and Yaling Yang

Department of Psychology and Neuroscience Program, University of Southern California, USA

A common feature of the antisocial, rule-breaking behavior that is central to criminal, violent and psychopathic individuals is the failure to follow moral guidelines. This review summarizes key findings from brain imaging research on both antisocial behavior and moral reasoning, and integrates these findings into a neural moral model of antisocial behavior. Key areas found to be functionally or structurally impaired in antisocial populations include dorsal and ventral regions of the prefrontal cortex (PFC), amygdala, hippocampus, angular gyrus, anterior cingulate and temporal cortex. Regions most commonly activated in moral judgment tasks consist of the polar/medial and ventral PFC, amygdala, angular gyrus and posterior cingulate. It is hypothesized that the rule-breaking behavior common to antisocial, violent and psychopathic individuals is in part due to impairments in some of the structures (dorsal and ventral PFC, amygdala and angular gyrus) subserving moral cognition and emotion. Impairments to the emotional component that comprises the feeling of what is moral is viewed as the primary deficit in antisocials, although some disruption to the cognitive and cognitive-emotional components of morality (particularly self-referential thinking and emotion regulation) cannot be ruled out. While this neurobiological predisposition is likely only one of several biosocial processes involved in the etiology of antisocial behavior, it raises significant moral issues for the legal system and neuroethics.

Keywords: antisocial; psychopathy; moral; prefrontal; temporal

INTRODUCTION

The burgeoning field of social neuroscience is beginning to provide important insights into the neural mechanisms that underlie the cognitive and affective processes that guide social behavior in everyday life. One particularly important sub-field within this area that has significant societal implications concerns the neural basis to antisocial behavior. The perspective that will be developed here is that there are some similarities between the neural system underlying moral decision-making in normal individuals, and brain mechanisms thought to be impaired in delinquent, criminal, violent and psychopathic populations. We suggest that this is not a chance association but instead represents a neural insight into the etiology of antisocial behavior.

A basic tenet underlying this thesis is that despite some differences in the constructs of childhood conduct disorder, antisocial personality disorder (APD), violence and psychopathy, a critically important common denominator to all is the failure to conform to the commonly accepted mores of society. Clearly, not all ‘immoral’ behavior is illegal, and not all features of a psychopathic personality pertain to immorality; there is more to antisocial behavior than a breakdown in the neural networks subserving moral thinking and feeling. Nevertheless, the overlap between morality and antisocial disorders is substantial, and it is

argued here that this is partly accounted for by disruption to neural systems common to both.

A general overview of studies of the neural structures impaired in those with persistent and significant antisocial behavior will first be outlined. The developing knowledge-base on neural mechanisms underpinning moral judgment will then be presented, with an emphasis on the best-replicated neural correlates of moral decision-making. Differences and similarities between moral and antisocial neural correlates will be delineated, and a hypothetical description provided of how disruption to both affective and cognitive components of morality may predispose to the rule-breaking that is central to antisocial disorders.

IMAGING FINDINGS ON ANTISOCIAL, VIOLENT AND PSYCHOPATHIC GROUPS

Despite the increasing evidence for neurological impairment in antisocial individuals, very few structural and functional brain imaging studies have been conducted specifically on the recognized medical disorder for antisocial behavior, i.e. APD. For this reason, the following review will include brain imaging studies on antisocial, violent and psychopathic populations as well as those with APD. While these groups make up different populations, the key thesis of this study is that a disruption to the neural systems underlying moral thinking and feeling gives rise to an ‘antisocial tendency’ (rule-breaking behavior), a key common facet to all antisocial conditions. We group the structural and functional imaging findings based

Received 11 August 2006; Accepted 17 September 2006

Correspondence should be addressed to Adrian Raine, Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061 USA. E-mail: raine@usc.edu.

on anatomical structures as follows: frontal lobe [orbitofrontal cortex (OFC)/dorsolateral prefrontal cortex (DLPFC)], temporal lobe (superior temporal gyrus/amygdala/hippocampus) and other brain areas (parietal lobe/angular gyrus/cingulate cortex).

Frontal lobe

Global prefrontal abnormalities. Prefrontal impairments are, perhaps, the best-replicated finding in the imaging literature on antisocial behavior. Reduced prefrontal glucose metabolism has been observed in murderers compared with normal controls (Raine *et al.*, 1994), while acts of aggressive impulsive behavior have been associated with reduced metabolism in the orbitofrontal, anterior medial frontal, and left anterior frontal regions (Goyer *et al.*, 1994). Using single photon emission computed tomography (SPECT), several studies have found significant correlations between reduced frontal blood flow and increased antisocial, aggressive behaviors (Oder *et al.*, 1992; Kuruoglu *et al.*, 1996; Gerra *et al.*, 1998; Soderstrom *et al.*, 2000, 2002).

These functional impairments are parallel by some evidence for structural prefrontal impairments. Three anatomical MRI (aMRI) studies have found significantly reduced prefrontal gray matter in antisocial and psychopathic individuals (Raine *et al.*, 2000; Yang *et al.*, 2005) and aggressive patients with temporal lobe epilepsy (Woermann *et al.*, 2000), although there is also one null finding (Dolan *et al.*, 2002). Similar findings were reported by Kruesi *et al.*, (2004) where a large (16%) reduction in prefrontal gray volume was found in children diagnosed with conduct disorder. However, given the small sample size ($N=10$), this finding was not statistically significant. Although these findings suggest widespread prefrontal deficits in APD, new evidence has begun to accumulate indicating that the abnormality in APD may be localized to the orbitofrontal and dorsolateral prefrontal regions.

Orbitofrontal cortex (OFC). Neurological research has indicated that individuals with orbitofrontal lesions are typically disinhibited, impulsive and unconcerned with the consequences of their behavior (Rolls *et al.*, 1994; Brower and Price, 2001). However, only two aMRI studies to date show reduced gray matter volumes in the OFC in antisocial individuals (Laakso *et al.*, 2002; Yang *et al.*, 2006a). Regarding function, two positron emission tomography (PET) studies have found reduced glucose metabolism in the OFC and medial PFC in impulsive patients and aggressive children (Juhasz *et al.*, 2001; Siever *et al.*, 1999). Similar findings were reported in another two PET studies showing significant negative correlations between aggression and metabolism rate in the OFC and medial frontal cortex (Goyer *et al.*, 1994; Pietrini *et al.*, 2000). In addition, several fMRI studies have shown abnormal OFC activation in impulsive individuals during response inhibition (Horn *et al.*, 2003) and in APD patients during both inhibitory

control (Vollm *et al.*, 2004) and fear-conditioning tasks (Birbaumer *et al.*, 2005).

Dorsolateral Prefrontal Cortex (DLPFC). Patient studies have shown that damage to the DLPFC classically leads to problems in planning, attention shifting, decision-making and perseverative responding (Manes *et al.*, 2002; Gomez-Beldarrain *et al.*, 2004). However, only one structural study to date has assessed prefrontal subregions, finding reduced gray matter volumes in the left DLPFC and the right OFC in alcoholics with antisocial personalities compared with controls (Laakso *et al.*, 2002). Regarding functional studies, reduced metabolism in the DLPFC has been found in aggressive patients (Hirono *et al.*, 2000) and also aggressive children with epilepsy (Juhasz *et al.*, 2001). In addition, two fMRI studies have observed abnormal DLPFC functioning in APD patients during both an emotional task (Schneider *et al.*, 2000) and also in an inhibition task (Vollm *et al.*, 2004).

Temporal lobe

The temporal lobe is the second major brain area traditionally associated with antisocial and aggressive behavior (Raine, 1993). Widespread abnormalities have been reported in several studies. Regarding structure, Kruesi *et al.* (2004) found a significant association between early-onset conduct disorder (without substance abuse comorbidity) and smaller temporal gray matter volumes. Similar findings of reduced temporal lobe volume have been reported in both incarcerated psychopaths (Dolan *et al.*, 2002) and in APDs (Barkataki *et al.*, 2006).

Regarding function, two SPECT studies have documented reduced temporal lobe functioning in aggressive patients (Volkow and Tancredi, 1987; Amen *et al.*, 1996). Another SPECT study found significant negative correlations between psychopathy and temporal perfusion, particularly left temporal blood flow (Soderstrom *et al.*, 2002). One fMRI study on violent offenders has shown reduced functioning in the temporal cortex compared with non-aggressive controls (Raine *et al.*, 2001). There are also some failures to observe temporal lobe functional impairments; one PET study of violent offenders failed to observe reduced temporal lobe glucose metabolism (Raine *et al.*, 1997a).

These findings are beginning to confirm global structural and functional temporal lobe abnormalities, although results are not entirely consistent. As now outlined, some studies have attempted to localize abnormalities to the middle, anterior inferior, superior and medial temporal (amygdala and hippocampus) regions.

Middle temporal gyrus. One PET study found reduced bilateral metabolism in the middle temporal gyrus in aggressive children with temporal lobe epilepsy (Juhasz *et al.*, 2001). A recent SPECT study also revealed reduced blood flow in the right middle temporal gyrus in APD patients (Goethals *et al.*, 2005).

Anterior inferior temporal cortex. Using PET, Wong *et al.*, (1997a) found reduced metabolism in the anterior inferior temporal cortex in two separate groups of violent patients. Similar findings were reported in a SPECT study suggesting a significant reduction in regional blood flow in the left anterior inferior temporal cortex (Hirono *et al.*, 2000).

Superior temporal gyrus. One study has revealed activation deficits in antisocial and psychopathic individuals localized to the *right* posterior superior temporal gyrus in a semantic processing task (Kiehl *et al.*, 2004). Conversely, a second study observed reduced activation in the *left* hemisphere superior temporal gyrus in a memory task (Kumari *et al.*, 2006).

Medial temporal cortex. In functional terms, two PET and one SPECT studies have found abnormal glucose metabolism in the medial temporal cortex. Volkow *et al.*, (1995) found significantly reduced metabolism in the medial temporal cortex in psychiatric patients with a history of repetitive violent behavior. The other study showed a similar metabolic reduction in the medial temporal lobes in violent offenders (Seidenwurm *et al.*, 1997). Using SPECT, Soderstrom *et al.*, (2000) found significant blood flow reductions in the right medial temporal region in impulsive violent criminals.

Amygdala. Three structural imaging studies have found significant reductions in the volume of the amygdala in violent offenders (Wong *et al.*, 1997b; Tiihonen *et al.*, 2000) and psychopathic individuals (Yang *et al.*, 2006b). Conversely, no volumetric difference have been found in aggressive patients with temporal lobe epilepsy (Trimble and Van Elst *et al.*, 1999; Van Elst *et al.*, 2000), although the latter study also reported a significantly higher rate of amygdala atrophy (20%) in aggressive patients.

In terms of functioning, abnormalities in the amygdala have been found in one PET study one magnetic resonance spectroscopy (MRS) study, and several fMRI studies. In the PET study, Raine *et al.* (1997a) found abnormal functional asymmetries in murderers, showing lower left and increased right amygdala functioning. Critchley *et al.* (2000) using MRS found violent patients with mild mental retardation to have reduced metabolism in the right amygdalo-hippocampal complex compared with non-violent controls. Somewhat surprisingly, two fMRI studies have reported *increased* amygdala activation in antisocial individuals while viewing negative visual content (Muller *et al.*, 2003) and during an aversive conditioning task (Schneider *et al.*, 2000). In contrast, *reduced* activation in the amygdala during the processing of affective stimuli has been found in criminal psychopaths (Kiehl *et al.*, 2001), psychopathic individuals (Veit *et al.*, 2002; Birbaumer *et al.*, 2005), and adolescents with conduct disorders (Sterzer *et al.*, 2005).

Hippocampus. Several imaging studies have examined the functional integrity of this region and found

abnormalities in murderers (Raine *et al.*, 1998), criminal psychopaths (Kiehl *et al.*, 2001) and violent offenders (Soderstrom *et al.*, 2000; Critchley *et al.*, 2000). With regard to structure, volumetric reductions in the hippocampus have been found in several aMRI studies of psychopaths (Raine *et al.*, 2004), violent offenders with APD (Laakso *et al.*, 2000), violent psychiatric patients (Kumari *et al.*, 2006; Barkataki *et al.*, 2006) and antisocial alcoholics (Laakso *et al.*, 2001).

Other Brain Areas

Several structural and functional imaging studies have suggested that the parietal lobe (particularly the angular gyrus) and anterior/posterior cingulate gyrus may also be compromised in antisocial groups.

Parietal Lobe. Regarding function, reduced metabolism has been found in the superior parietal cortex in aggressive patients (Hirono *et al.*, 2000), murderers (Raine *et al.*, 1997a) and individuals with impulsive personality disorders (Siever *et al.*, 1999). However, no structural MRI study has been conducted to date showing structural impairments in individuals with APD.

Only two studies to our knowledge have assessed functioning in the angular gyrus in individuals with APD. Both found significantly reduced activation, but in different hemispheres. Raine *et al.*, (1997a) in a PET study found that murderers have reduced glucose metabolism in the *left* angular gyrus. Using SPECT, Soderstrom *et al.*, (2000) found a significant blood flow reduction in the *right* angular gyrus in impulsive violent criminals.

Anterior and Posterior Cingulate. Several fMRI studies have provided evidence showing functional impairments in the anterior and posterior cingulate cortex (PCC) in APDs. Five fMRI studies have shown reduced activation in the anterior cingulate cortex (ACC) in criminal psychopaths during an affective memory task (Kiehl *et al.*, 2001), APD patients during a working memory task (Kumari *et al.*, 2006), psychopaths during a fear-conditioning task (Birbaumer *et al.*, 2005), and conduct disorder patients during viewing of negative emotional pictures (Sterzer *et al.*, 2005). In addition, an abnormal activation pattern was reported in APD patients during a response inhibition task (Vollm *et al.*, 2004). With regard to the PCC, two functional imaging studies have also found reduced activation in this region in criminal psychopaths (Kiehl *et al.*, 2001) and aggressive patients (New *et al.*, 2002).

Summary and interpretation of imaging findings in antisocial populations

A fairly sizable imaging literature has now been built up on functional and (to a lesser extent) structural brain abnormalities in diverse antisocial groups, although for some structures such as the amygdala and hippocampus, the precise pattern of results is somewhat complex. Taken together, the best-replicated brain imaging abnormality

found to date across a wide variety of antisocial groups, across structure and function and across different imaging methodologies is the PFC. While dysfunction in the OFC may be associated with poor inhibitory control, emotional decision-making and reward/punishment processing in antisocial individuals, the additional involvement of DLPFC dysfunction may also predispose to response perseveration (life-long 'revolving door' antisocial behavior despite repeated punishment), and poor planning/organization (resulting in an occupationally and socially dysfunctional lifestyle).

In addition to the PFC, there is increasing evidence for structure/function impairments in the amygdala (poor fear conditioning), hippocampus (emotion regulation and contextual fear conditioning), temporal cortex (language and memory), anterior cingulate (autonomic functions and emotion regulation) and initial evidence for reduced functioning in the angular gyrus (reading and arithmetic—potentially predisposing to school/occupational failure). Impairments to varying degrees have been documented in antisocial populations in all these neurocognitive processes. It is clear that multiple brain regions may contribute to antisocial behavior for different reasons. Furthermore, future research needs to move beyond the simple identification of single structures impaired in antisocials to the more complex delineation of the neural circuits that are impaired. In all likelihood, no single brain structure is critical in the development of antisocial behavior. Instead, the greater the number of neural impairments across different cognitive and affective domains related to an antisocial lifestyle, the higher the likelihood of an antisocial outcome. The remainder of this article outlines how impairments to the circuitry underlying morality may be one such important additional risk factor for the antisocial spectrum of disorders.

NEUROLOGICAL STUDIES OF ANTISOCIAL BEHAVIOR

Neurological studies of patients suffering trauma to discrete brain regions have provided invaluable data on neural mechanisms predisposing to antisocial behavior. An inevitable limitation of imaging findings is that they are correlational, while, in contrast, the natural accidents that comprise neurological findings bear more directly on causality. Intriguingly, findings from these studies converge with evidence from brain imaging studies, particularly with respect to the PFC.

Two key prefrontal neurological syndromes that bear on antisocial behavior have been delineated on the basis of the timing of neurological damage, and give rise to slightly different antisocial outcomes. 'Acquired sociopathy', in which accidental damage occurs to the ventromedial PFC in adulthood, has been shown to result in pseudo-psychopathic, disinhibited, antisocial behavior (Damasio *et al.*, 1990; Damasio, 1994), together with bad decision-making and reduced anticipatory skin conductance

responses to stimuli predicting negative outcomes. Nevertheless, with some exceptions (e.g. Blair and Cipollotti, 2000) these patients do not generally exhibit impairments in moral judgment. For example, patient E.V.R. suffered damage to the ventromedial PFC at the age of 35 years, and consequently manifested significant antisocial behavior and reduced skin conductance responsivity to socially meaningful stimuli (Damasio *et al.*, 1990), yet was unimpaired on a moral reasoning task and showed good abstract social insight (Saver and Damasio, 1991).

In the second syndrome of developmental sociopathy, prefrontal damage early in childhood is similarly associated with a history of antisocial and aggressive behavior together with reduced anticipatory skin conductance responses to stimuli associated with punishment that extends throughout childhood and into adulthood. Unlike acquired sociopathy, however, developmental sociopathy is associated with significant impairments in moral reasoning and judgment. For example, two case studies reported by Anderson *et al.*, (1999) show that lesions to either the polar or ventromedial PFC occurring in the first 16 months of life result in lifelong psychopathic-like antisocial behavior, and also impaired social and moral reasoning as assessed by Kohlberg's moral judgment task. These studies would suggest, therefore, that the PFC, particularly polar and ventromedial sectors, are critically involved in the development (but not necessarily retention) of moral reasoning.

Brain imaging and neurological findings on antisocial behavior show convergence in other ways. The somatic marker hypothesis argues that somatic states associated with previous stimuli/events provide an effective, automatic mechanism to facilitate choice of response options (Damasio *et al.*, 1990; Damasio, 1994). It is further argued that damage to the ventromedial PFC effectively deactivates this mechanism, resulting in poor decision-making and antisocial behavior (Damasio *et al.*, 1990; Bechara *et al.*, 1997). These findings from neurological patients gain some support from the imaging literature. Individuals with APD have been found by MRI to have both a reduced volume of the PFC together with reduced skin conductance during a social stressor that elicits somatic marker states of embarrassment, shame and guilt (Raine *et al.*, 2000). Furthermore, the antisocials with the greatest prefrontal volume reduction show the greatest autonomic impairment. Different clinical neuroscience paradigms consequently converge on a common conclusion that prefrontal cortical impairment is a significant risk factor for the development of antisocial behavior.

NEURAL BASIS TO MORAL JUDGMENTS

Behavior that breaks the moral guidelines set down by society is a fundamental feature of antisocial disorders, and almost defines criminal behavior. In this context,

recent imaging research is beginning to identify which cortical areas are activated when subjects perform tasks involving moral conundrums. This section provides a brief overview of what we have learnt so far about brain mechanisms subserving moral decision-making from 12 fMRI studies. As will be seen, a wide variety of tasks have been used asking somewhat different questions on the neural correlates of morality. Despite this diversity, a number of key brain areas appear to be a common denominator for moral information-processing.

Regions consistently activated—medial PFC, angular gyrus and ventral PFC

The ground-breaking study in this area (Greene *et al.*, 2001) focused on the specific difference between making judgments (i.e. 'appropriate' or 'inappropriate') on 'moral personal' dilemmas (e.g. throwing a person out of a sinking life-boat to save others), and 'moral impersonal' dilemmas (e.g. keeping money found in a lost wallet). Moral dilemmas involving a personal component, compared with both impersonal moral dilemmas and nonmoral dilemmas, activated the medial frontal gyrus (BA 9 and 10), the posterior cingulate (BA 31) and both left and right angular gyri (BA 39). This initial study suggested that these structures play a central role in the emotional processes that influence personal moral decision-making.

Studies since 2001 have confirmed the importance of the medial PFC, angular gyrus and posterior cingulate in processing moral stimuli, and at the same time implicated a further structure—the ventral PFC. The medial PFC (and the frontal pole, BA 10) is activated by passive viewing of pictures depicting moral *vs* nonmoral violations (Moll *et al.*, 2002a; Harenski and Hamann, 2006), making judgments on auditory moral *vs* nonmoral sentences (Oliveira-Souza and Moll, 2000), passive viewing of morally disgusting *vs* nonmorally disgusting statements (Moll *et al.*, 2005); difficult *vs* easy moral dilemmas (Greene *et al.*, 2004), moral decision-making *vs* semantic decision-making (Heekeren *et al.*, 2003), personal *vs* impersonal moral dilemmas (Green *et al.*, 2004), judgment on moral *vs* non-moral actions (Borg *et al.*, 2006) and 'utilitarian' moral decision-making (e.g. acceptance of loss of life for the greater good) *vs* 'nonutilitarian' decision-making (prohibiting a loss of life even though more lives could be saved—Greene *et al.*, 2004). Overall, with only a few exceptions, studies consistently observe activation of fronto-polar cortex (particularly the medial aspects of BA 10) in moral judgment tasks.

The angular gyrus (BA 39—in some studies labeled as posterior superior temporal gyrus) lying in the parietal lobe at the intersection of temporal, parietal and occipital cortices, surrounds the tip of the posterior superior temporal sulcus and is a second region that shows robust activation across different moral judgment tasks. Angular gyral activation initially observed by Greene *et al.*, (2001)

has also been observed in passive viewing of pictures depicting moral *vs* nonmoral violations (Moll *et al.*, 2002a; Harenski and Hamann 2006), making judgments on auditory moral *vs* nonmoral sentences (Oliveira *et al.*, 2000), moral *vs* semantic decision-making (Heekeren *et al.*, 2003; replicated in Heekeren *et al.*, 2005), personal *vs* impersonal moral dilemmas (Greene *et al.*, 2004), judgments on moral *vs* nonmoral actions (Borg *et al.*, 2006) and difficult *vs* easy moral judgments (Greene *et al.*, 2004). In addition, Moll *et al.*, (2002b) found activation in the area of the superior temporal sulcus bordering the angular gyrus when responding to unpleasant moral *vs* unpleasant non-moral statements. The weight of support for angular gyral involvement in moral tasks, in association with the posterior superior temporal sulcus, is consequently as strong as that for medial PFC involvement.

One additional brain region that could not be imaged in the initial study by Greene *et al.*, (2001) due to susceptibility artifact, yet which is being increasingly implicated in moral judgment tasks, is the ventral PFC. This region encompasses the OFC and the gyrus rectus (also broadly termed ventrolateral and ventromedial, respectively). Activation in this region has been found during passive viewing of pictures depicting moral *vs* non-moral violations (Moll *et al.*, 2002a—right orbitofrontal), responding to unpleasant moral *vs* unpleasant nonmoral statements (Moll *et al.*, 2002b—gyrus rectus and orbitofrontal), passive viewing of morally disgusting *vs* non-morally disgusting statements (Moll *et al.*, 2005—orbitofrontal bilaterally), judgments on moral *vs* non-moral actions (Borg *et al.*, 2006—orbitofrontal), 'automatic' moral judgment to high *vs* low immoral stimuli (Luo *et al.*, 2006—medial orbitofrontal) and moral *vs* semantic decision-making (Heekeren *et al.*, 2003—ventromedial; replicated in Heekeren *et al.*, 2005).

Regions less consistently activated—posterior cingulate, amygdala and temporal pole

While the polar/medial PFC, ventral PFC and angular gyrus encapsulate areas with the strongest evidence for activation during moral tasks, three other regions also need to be considered. Activation of the posterior cingulate has been observed during moral personal *vs* impersonal dilemmas (Greene *et al.*, 2001), moral *vs* semantic decision-making (Heekeren *et al.*, 2005) and passive viewing of pictures depicting moral *vs* non-moral violations (Moll *et al.*, 2002a; Harenski and Hamann 2006). In addition, Green *et al.*, (2004) observed consistent activation of the posterior cingulate across all three of their experimental moral decision-making conditions/comparisons (personal *vs* impersonal, difficult *vs* easy and utilitarian *vs* non utilitarian). Several studies have also reported amygdala activation (Moll *et al.*, 2002a; Greene *et al.*, 2004; Berthoz *et al.*, 2006; Harenski and Hamann 2006; Luo *et al.*, 2006). The temporal pole has also been activated in some

studies (Oliveira-Souza *et al.*, 2000; Moll *et al.*, 2002b; Heekeren *et al.*, 2003; Heekeren *et al.*, 2005). In contrast, two areas that are centrally associated with cognitive-affective processing yet which have (with notable exceptions) not been activated in most studies are the anterior cingulate (Berthoz *et al.*, 2006; Greene *et al.*, 2004) and the insula (Moll 2002a; Greene *et al.*, 2004).

A notable feature of the replicated activation findings for the polar/medial PFC, ventromedial/orbital PFC, angular gyrus, posterior cingulate and to some extent the amygdala in moral tasks is that studies differ quite widely on viewing conditions (active *vs* passive), task requirements, and emotion/cognitive control tasks. The fact that activation is found across diverse conditions attests to the robustness of findings for these particular structures. This in turn suggests that they form a core of the moral neural circuit, although other structures may play more limited roles in certain cases of moral decision-making. As one example, the fact that the anterior cingulate is activated during difficult compared to easy moral decision-making has been interpreted as consistent with the involvement of this area in cognitive conflict (Greene *et al.*, 2004).

The regulation of moral emotions

The above studies provide important clues as to the neural circuit underlying moral emotions and cognitions. A different question that can be asked concerns which brain areas are involved in the *regulation* of moral emotions. In addressing this issue, Harenski and Hamann (2006) not only assessed activation during unpleasant moral *vs* unpleasant nonmoral social picture viewing, but also requested participants to suppress their emotional reactions to both sets of stimuli. As expected, the two brain areas activated by moral (*vs* nonmoral) conditions were the angular gyrus and the posterior cingulate. In contrast, suppression of emotion to pictures of moral transgressions leads to a reduction in amygdala activation and an increase in polar/medial PFC (area 10) activation. These findings suggest a unique activation role for area 10 of the PFC in both the process of moral decision-making and in the regulation of moral emotions, while deactivation of the amygdala during down-regulation of moral emotions provides further credence to its active involvement in the generation of moral emotions.

MORAL CIRCUITRY AND ANTISOCIAL BEHAVIOR IN AN EVOLUTIONARY CONTEXT

It is becoming increasingly clear that moral judgments are not the sole product of thoughtful introspection conducted in isolation of emotion. In contrast, the imaging literature would suggest that brain systems underlying affective as well as cognitive processes are active when individuals weigh their actions within a moral context.

Furthermore, there is likely a complex interplay between cognitions and emotions during the formulation of a moral decision, blurring the distinction between thinking and feeling, at least in modern-day man.

It is likely that some (if not many) of our moral values are shrouded in a deep evolutionary history where emotions—not cognitions—constituted the driving force of moral action. The astonishing success of early hominids was heavily predicated on reciprocal altruism and a resource-sharing social structure. Nevertheless, selfishness (taking but not giving resources) can constitute a competing evolutionary stable strategy that has to be kept in check for the survival of the species. Negative moral emotions likely evolved to counteract the breaking of social conventions. Moral feelings of indignation, disdain, disgust and contempt can give rise to the stronger emotions of outrage and vengeance that then give rise to ostracization of the cheat from the social group, injury or even death. At this level, morality is largely emotion-driven, relatively automatic, and has little or no higher cognitive control component in early hominids. As hominid society became more complex, higher-order cognitive processes likely became increasingly important for both dealing with more complex moral dilemmas, and for regulating the expression of moral emotions.

Despite the evolution of social mechanisms to deter antisocial ‘cheating’ behavior, it has been argued that antisocial behavior is an evolutionary stable strategy—a pre-programmed behavioral approach that maximizes reproductive fitness. Psychopathy has been viewed as the full expression of this ‘cheating’ strategy (Raine, 1993). At low base rates within the population, psychopaths can be successful in extracting resources from other individuals before moving on to other social groups to avoid the consequences of moralistic rage and retributive justice. Psychopathic traits of superficial charm, egocentrism, manipulation, pathological lying and deception, promiscuous sexual behavior, lack of remorse and guilt, superficial relationships and general parasitic lifestyle are viewed as key components of this evolutionary stable strategy that are observable at a surface level (Raine, 1993). Importantly, psychopaths adopt a transient, unstable, ‘stimulation-seeking’ lifestyle, moving from one place to another to avoid ultimate detection—an essential antidote to the moral medicine that would otherwise be metered out to them.

In this evolutionary context, an essential component of this successful cheating strategy *must* be a gene machine that lacks a core moral sense. One way to create such a cheating machine would be to engineer individuals lacking the neural circuitry essential for moral feelings and behavior. One prediction generated by this model is that antisocial and psychopathic individuals would manifest impairments in the brain mechanism requisite for morals, particularly those neural processes critical to the

experiencing of moral emotions—the brakes on rule-breaking behavior.

NEURAL MODELS OF MORALITY AND ANTISOCIALITY

The review outlined above concludes that the brain regions most implicated to date in antisocial, aggressive and psychopathic behavior are the prefrontal (dorsolateral and ventral) and temporal cortices. Some evidence also exists implicating the amygdala, hippocampus, angular gyrus and anterior cingulate. It has also been argued that the best-replicated neural correlates of morality are the polar/medial PFC, ventral PFC and angular gyrus, with significant support also for the posterior cingulate and amygdala. We also argue that a critical driving force of antisocial and psychopathic behavior is a disruption to the neural circuitry underlying moral thinking and feeling. Figure 1 juxtaposes these two sets of empirical data to create an initial neural model of morality and antisociality, highlighting those regions impaired *only* in antisocial populations (red), areas activated *only* by moral judgment tasks (green) and areas common to *both* antisociality and morality (yellow).

It can be seen both that there are substantial areas of overlap between antisocial/psychopathic behavior and moral judgment-emotion, and also significant differences. Brain regions common to both include ventral and polar/medial PFC sectors, the amygdala and angular gyrus/posterior superior temporal gyrus. A key difference is that while there is increasing evidence for hippocampal and anterior cingulate impairment in antisocial/psychopathic individuals, moral studies consistently fail to observe selective activation in these structures. Similarly, there is some replicable evidence implicating the posterior cingulate in moral judgment tasks, but evidence implicating this region in antisocial behavior is to date sparse. A further qualification is that only two studies have documented reduced functioning in the angular gyrus in antisocial populations, although few studies have explored this parietal structure which should be a target region-of-interest in future imaging studies.

The partial overlap of structures implicated in antisocial populations and moral judgment tasks gives rise to the hypothesis that *some* of the brain impairments found in antisocial individuals disrupt moral emotion/decision-making, which in turn predisposes the individual to rule-breaking, antisocial behavior. A critical question raised by this theoretical perspective of antisocial behavior concerns which component of ‘morality’ is impaired—cognition (i.e. moral ‘reasoning’), emotion (i.e. moral ‘feelings’) or both?

A tentative answer would be that the emotional component is most impaired in antisocial, psychopathic populations. Regarding basic cognitive processes involved in moral decision-making, at a fundamental level there is

little question that almost all criminal and psychopathic individuals know right from wrong. While some evidence exists for a difference in level of moral reasoning in delinquent, criminal and psychopathic groups (Raine, 1993; Blair, 1995), antisocial behavior could cause differences in moral thinking, rather than vice versa. That is, living an antisocial way of life may change moral thinking to justify the individual’s repeated antisocial actions and reduce cognitive dissonance. Furthermore, it has also been argued that psychopaths show excellent (not poor) moral reasoning ability when discussing hypothetical situations—their real failure comes in applying their excellent moral conceptual formulations to guiding their own behavior (Cleckley, 1976).

The feeling of what's right and wrong

Given the above, it is suggested that it is predominantly the *feeling* of what is moral that is deficient in antisocial groups, rather than the *knowing* of what is moral. This moral feeling, centered on the PFC and amygdala, is the engine that translates the cognitive recognition that an act is immoral into behavioral inhibition—and it is this engine that functions less well in antisocial, violent and psychopathic individuals.

We also hypothesize that while deficits in the affective component of morality is the primary impairment in antisocial individuals, cognitive components of morality could also be compromised for three reasons. First, both the polar/medial PFC and posterior cingulate have been shown to play an important role in self-appraisal and self-reflection (Ochsner *et al.*, 2005; Johnson *et al.*, 2006), while the OFC is also involved in the process of self-perception and insight (Beer *et al.*, 2006). If an individual is unable to relate back onto themselves the negative emotion associated with the thought of perpetrating an immoral act due to impairments to either the medial PFC or posterior cingulate, they may become predisposed to rule-breaking behavior, despite intact emotional processing. Second, cognition and emotion in moral decision-making cannot be easily dissociated. The medial PFC is activated during the suppression of moral emotions (Harenski and Hamann 2006). If an individual lacks this regulatory control, there may be an inability to down-regulate moral outrage triggered by a third party’s negative acts, resulting in impulsive, reactive aggression by that individual, and/or more planned, controlled retaliatory actions. Third, and relatedly, the angular gyrus has been speculated to be associated with a sense of responsibility for one’s actions (Borg *et al.*, 2006); while intact moral emotions may normally place a brake on rule-breaking behavior, the lack of a sense of responsibility may move the individual more in an immoral direction if the rewards are sufficient. Consequently, both emotional and cognitive components of the morality

circuit may be implicated in antisocial, psychopathic behavior.

CONCLUSIONS AND SUMMARY

In summary, we have argued the following:

(i) brain regions compromised in antisocial, violent and psychopathic populations include both dorsal and ventral regions of the PFC, amygdala, hippocampus, angular gyrus, anterior cingulate and temporal cortex including the superior temporal gyrus;

- (ii) regions activated during moral decision-making in normals include the polar/medial PFC, ventral PFC, angular gyrus, amygdala and posterior cingulate;
- (iii) brain areas associated with both moral reasoning and antisocial behavior significantly overlap;
- (iv) the rule-breaking, immoral behavior of antisocial and psychopathic individuals may in part be due to impairments in those brain regions subserving moral cognition and emotion;
- (v) while impairments to the moral emotional system may be primary in antisocials, disruption of moral cognitive and cognitive-emotional systems are also possible.

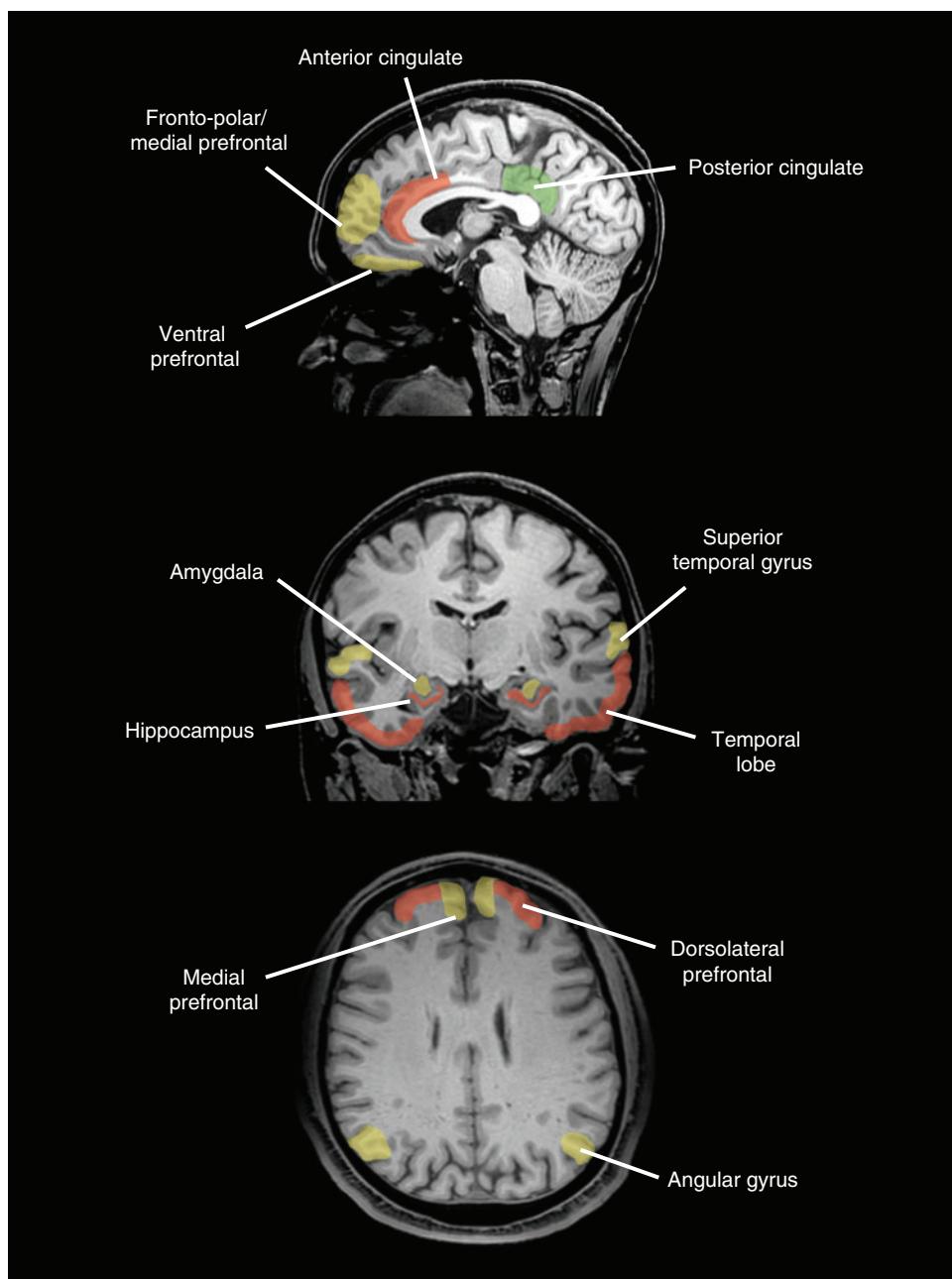


Fig. 1 A schematic diagram of brain regions impaired only in antisocial groups (red), activated only in moral decision-making (green) and regions common to both antisocial behavior and moral decision-making (yellow).

This neuro-moral theory of antisocial, violent and psychopathic behavior must be regarded as provisional. The precise foci of the likely multiple neural deficits in antisocial groups (particularly within the large regions compromising prefrontal and temporal cortex) remains to be delineated, while an understanding of the neural basis to moral decision-making is clearly still in its infancy. For example, there are hints that temporal pole impairment/activation may be implicated in both antisociality/morality, respectively, yet, limitations in the evidence to date precluded inclusion in Figure 1. A compromised moral circuit will be only one of multiple etiological processes ultimately found to predispose to the complex constructs of antisocial/aggressive/psychopathic behavior. Interactions between neural and social risk factors for antisocial behavior cannot be ignored (Raine *et al.*, 1997b; Caspi *et al.*, 2002), raising yet another layer of complexity to a full elucidation of antisocial disorders.

Despite these caveats, we believe that a neural moral hypothesis of antisocial disorders is worthy of examination in future studies. Neuroscientists and lawyers alike are beginning to raise important questions about the implications of new neuroscience knowledge for society, the law, and civil liberties (Morse, 2004), leading to the beginning of a new sub-discipline of 'neuroethics' (Farah, 2004). Psychopaths may not be 'morally insane' in any strict legal sense as they are cognitively capable of distinguishing right from wrong, but if they lack the capacity for the *feeling* of what is moral due to neurobiological impairments beyond their control, are they *fully* responsible for their criminal behavior? If not, what are the implications for punishment and our concepts of both justice and retribution? This challenging question that lies at the interface of law, neuroscience and neuroethics, begs for further enlightenment from future systematic imaging research on both morality-processing and antisocial behavior.

Conflict of Interest

None declared.

REFERENCES

- Amen, D.G., Stubblefield, M., Carmichael, B., Thisted, R. (1996). Brain SPECT findings and aggressiveness. *Annals of Clinical Psychiatry*, 8, 129–37.
- Anderson, S.W., Bechara, A., Damasio, H., Tranel, D., Damasio, A.R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2, 1032–7.
- Barkataki, I., Kumari, V., Das, M., Taylor, P., Sharma, T. (2006). Volumetric structural brain abnormalities in men with schizophrenia or antisocial personality disorder. *Behavioral Brain Research*, 15, 239–47.
- Beer, J.S., John, O.P., Scabini, D., Knight, R.T. (2006). Orbitofrontal cortex and social behavior: integrating self-monitoring and emotion-cognition interactions. *Journal of Cognitive Neuroscience*, 18, 871–9.
- Bechara, A., Damasio, H., Tranel, D., Damasio, A.R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science*, 275, 1293–4.
- Berthoz, S., Grezes, J., Armony, J.L., Passingham, R.E., Dolan, R.J. (2006). Affective response to one's own moral violations. *Neuroimage*, 31, 945–50.
- Birbaumer, N., Viet, R., Lotze, M., et al. (2005). Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study. *Archive of General Psychiatry*, 62, 799–805.
- Blair, R.J.R. (1995). A cognitive developmental approach to morality – investigating the psychopath. *Cognition*, 57, 1–29.
- Blair, R.J.R., Cipolotti, L. (2000). Impaired social response reversal. A case of "acquired sociopathy". *Brain*, 123, 1122–41.
- Borg, J.S., Hynes, C., Van Horn, J., Grafton, S., Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: an fMRI investigation. *Journal of Cognitive Neuroscience*, 18, 803–17.
- Brower, M.C., Price, B.H. (2001). Advances in neuropsychiatry: neuro-psychiatry of frontal lobe dysfunction in violent and criminal behaviour: a critical review. *Journal of Neurology, Neurosurgery, and Psychiatry*, 71, 720–726.
- Caspi, A., McClay, J., Moffitt, T., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297, 851–4.
- Cleckley, H. (1976). *The mask of sanity* 5th edn, St Louis: Mosby.
- Critchley, H.D., Simmons, A., Daly, E.M., et al. (2000). Prefrontal and medial temporal correlates of repetitive violence to self and others. *Biological Psychiatry*, 45, 928–34.
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.
- Damasio, A.R., Tranel, D., Damasio, H. (1990). Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. *Behavioural Brain Research*, 41, 81–94.
- Farah, M.J. (2004). Neuroethics: a guide for the perplexed. *Cerebrum*, 6, 29–38.
- Dolan, M.C., Deakin, J.F.W., Roberts, N., Anderson, I.M. (2002). Quantitative frontal and temporal structural MRI studies in personality-disordered offenders and control subjects. *Psychiatry Research Neuroimaging*, 116, 133–49.
- Gerra, G., Calbiani, B., Zaimovic, A., et al. (1998). Regional cerebral blood flow and comorbid diagnosis in abstinent opioid addicts. *Psychiatry Research*, 26, 117–26.
- Goethals, I., Audenaert, K., Jacobs, F., et al. (2005). Brain perfusion SPECT in impulsivity-related personality disorders. *Behavioral Brain Research*, 157, 187–92.
- Gomez-Beldarrain, M., Harries, C., Garcia-Monco, J.C., Ballus, E., Grafman, J. (2004). Patients with right frontal lesions are unable to assess and use advice to make predictive judgments. *Journal of Cognitive Neuroscience*, 16, 74–89.
- Goyer, P.F., Andreason, P.J., Semple, W.E., et al. (1994). Positron-emission tomography and personality disorders. *Neuropsychopharmacology*, 10, 21–8.
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44, 389–400.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293, 2105–8.
- Harenski, C.L., Hamann, S. (2006). Neural correlates of regulating negative emotions related to moral violations. *Neuroimage*, 30, 313–24.
- Hecker, H.R., Wartenburger, I., Schmidt, H., Prehn, K., Schwintowski, H.P., Villringer, A. (2005). Influence of bodily harm on neural correlates of semantic and moral decision-making. *Neuroimage*, 24, 887–97.
- Hecker, H.R., Wartenburger, I., Schmidt, H., Schwintowski, H.P., Villringer, A. (2003). An fMRI study of simple ethical decision-making. *Neuroreport*, 14, 1215–9.
- Hirono, N., Mega, M.S., Dinov, I.D., Mishkin, F., Cummings, J.L. (2000). Left frontotemporal hypoperfusion is associated with aggression in patients with dementia. *Archive of Neurology*, 57, 861–6.

- Horn, N.R., Dolan, M., Elliott, R., Deakin, J.F.W., Woodruff, P.W.R. (2003). Response inhibition and impulsivity: an fMRI study. *Neuropsychologia*, 41, 1959–66.
- Johnson, M.K., Raye, C.R., Mitchell, K.J., Touryan, S.R., Greene, E.J., Nolen-Hoeksema, S. (2006). Dissociating medial frontal and posterior cingulate activity during self-reflection. *Social, Cognitive, and Affective Neuroscience*, 1, 64.
- Juhasz, C., Behen, M.E., Muzik, O., Chugani, D.C., Chugani, H.T. (2001). Bilateral medial prefrontal and temporal neocortical hypometabolism in children with epilepsy and aggression. *Epilepsia*, 42, 991–1001.
- Kiehl, K.A., Smith, A.M., Hare, R.D., Mendrek, A., Forster, B.B., Brink, J. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry*, 50, 677–84.
- Kiehl, K.A., Smith, A.M., Mendrek, A., Forster, B.B., Hare, R.D., Liddle, P.F. (2004). Temporal lobe abnormalities in semantic processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Psychiatry Research Neuroimaging*, 130, 27–42.
- Kruesi, M.J.P., Casanova, M.F., Mannheim, G., Johnson-Bilder, A. (2004). Reduced temporal lobe volume in early onset conduct disorder. *Psychiatry Research Neuroimaging*, 132, 1–11.
- Kumari, V., Das, M., Hodgins, S., et al. (2006). Association between violent behavior and impaired prepulse inhibition of the startle response in antisocial personality disorder and schizophrenia. *Behavioral Brain Research*, 158, 159–66.
- Kuruoglu, A.C., Arikan, Z., Vural, G., Karatas, M. (1996). Single photon emission computerised tomography in chronic alcoholism: antisocial personality disorder may be associated with decreased frontal perfusion. *British Journal of Psychiatry*, 169, 348–54.
- Laakso, M.P., Gunning-Dixon, F., Vaurio, O., Repo-Tiihonen, E., Soininen, H., Tiihonen, J. (2002). Prefrontal volumes in habitually violent subjects with antisocial personality disorder and type 2 alcoholism. *Psychiatry Research: Neuroimaging*, 114, 95–102.
- Laakso, M.P., Vaurio, O., Koivisto, E., Savolainen, L., Eronen, M., Aronen, H.J. (2001). Psychopathy and the posterior hippocampus. *Behavioural Brain Research*, 118, 187–93.
- Luo, Q.A., Nakic, M., Wheatley, T., Richell, R., Martin, A., Blair, R.J.R. (2006). The neural basis of implicit moral attitude – An IAT study using event-related fMRI. *Neuroimage*, 30, 1449–57.
- Manes, F., Sahakian, B., Clark, L., et al. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, 125, 624–39.
- Moll, J., Oliveira-Souza, R., Bramati, I. E., Grafman, J. (2002a). Functional networks in emotional moral and nonmoral social judgments. *Neuroimage*, 16, 696–703.
- Moll, J., Oliveira-Souza, R., Eslinger, P.J., et al. (2002b). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *Journal of Neuroscience*, 22, 2730–6.
- Moll, J., Oliveira-Souza, R., Moll, F. T., et al. (2005). The moral affiliations of disgust – A functional MRI study. *Cognitive and Behavioral Neurology*, 18, 68–78.
- Morse, S. J. (2004). New neuroscience, old problems: legal implications of brain science. *Cerebrum*, 6, 81–90.
- Muller, J.L., Sommer, M., Wagner, V., et al. (2003). Abnormalities in emotion processing within cortical and subcortical regions in criminal psychopaths: evidence from a functional magnetic resonance imaging study using pictures with emotional content. *Psychiatry Research: Neuroimaging*, 54, 152–162.
- New, A.S., Hazlett, E.A., Buchsbaum, M.S., et al. (2002). Blunted prefrontal cortical 18fluorodeoxyglucose positron emission tomography response to meta-chlorophenylpiperazine in impulsive aggression. *Archives of General Psychiatry*, 59, 621–9.
- Ochsner, K.N., Beer, J.S., Robertson, E.R., et al. (2005). The neural correlates of direct and reflected self-knowledge. *Neuroimage*, 28, 797–814.
- Oder, W., Goldenberg, G., Spatt, J., Podreka, I., Binder, H., Deecke, L. (1992). Behavioural and psychosocial and regional cerebral blood flow: a SPECT study. *Journal of Neurology, Neurosurgery, and Psychiatry*, 55, 475–80.
- Oliveira-Souza, R., Moll, J. (2000). The moral brain: functional MRI correlates of moral judgment in normal adults. *Neurology*, 54, 252.
- Pietrini, P., Guazzelli, M., Bassi, G., Jaffe, K., Grafman, J. (2000). Neural correlates of imaginal aggressive behavior assessed by positron emission tomography in healthy subject. *American Journal of Psychiatry*, 157, 1772–81.
- Raine, A. (1993). *The psychopathology of crime: Criminal behavior as a clinical disorder*. San Diego: Academic Press.
- Raine, A., Brennan, P.A., Farrington, D.P., Mednick, S.A. (1997b). *Biosocial bases of violence*. New York: Plenum.
- Raine, A., Buchsbaum, M., LaCasse, L. (1997a). Brain abnormalities in murderers indicated by positron emission tomography. *Biological Psychiatry*, 42, 495–508.
- Raine, A., Ishikawa, S.S., Arce, E., et al. (2004). Hippocampal structural asymmetry in unsuccessful psychopaths. *Biological Psychiatry*, 55, 185–91.
- Raine, A., Lencz, T., Bihrl, S., LaCasse, L., Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry*, 57, 119–27.
- Raine, A., Meloy, J.R., Bihrl, S., Stoddard, J., LaCasse, L., Bushsbaum, M.S. (1998). Reduced prefrontal and increased subcortical brain functioning assessed using positron emission tomography in predatory and affective murderers. *Behavioral Science and the Law*, 16, 319–32.
- Raine, A., Park, S., Lencz, T., et al. (2001). Reduced right hemisphere activation in severely abused violent offenders during a working memory task: An fMRI study. *Aggressive Behavior*, 27, 111–129.
- Rolls, E.T., Hornak, J., Wade, D., McGrath, J. (1994). Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. *Journal of Neurology, Neurosurgery, and Psychiatry*, 57, 1518–24.
- Saver, J.L., Damasio, A.R. (1991). Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia*, 29, 1241–9.
- Schneider, F., Habel, U., Kessler, C., Posse, S., Grodd, W., Muller-Gartner, H.W. (2000). Functional imaging of conditioned aversive emotional responses in antisocial personality disorder. *Neuropsychobiology*, 42, 192–201.
- Seidenwurm, D., Pounds, T.R., Globus, A., Valk, P.E. (1997). Abnormal temporal lobe metabolism in violent subjects: correlation of imaging and neuropsychiatric findings. *American Journal of Neuroradiology*, 18, 625–31.
- Siever, L.J., Buchsbaum, M.S., New, A.S., et al. (1999). d,l-fenfluramine response in impulsive personality disorder assessed with [18F]fluorodeoxyglucose positron emission tomography. *Neuropsychopharmacology*, 20, 413–23.
- Soderstrom, H., Hultin, L., Tullberg, M., Wikkelso, C., Ekholm, S., Forsman, A. (2002). Reduced frontotemporal perfusion in psychopathic personality. *Psychiatry Research Neuroimaging*, 114, 81–94.
- Soderstrom, H., Tullberg, M., Wikkelso, C., Ekholm, S., Forsman, A. (2000). Reduced regional cerebral blood flow in non-psychotic violent offenders. *Psychiatry Research: Neuroimaging*, 98, 29–41.
- Sterzer, P., Stadler, C., Krebs, A., Kleinschmidt, A., Poustka, F. (2005). Abnormal neural responses to emotional visual stimuli in adolescents with conduct disorder. *Biological Psychiatry*, 57, 7–15.
- Tiihonen, J., Hodgins, S., Vaurio, O. (2000). Amygdaloid volume loss in psychopathy. *Society for Neuroscience Abstracts*, 2017.
- Trimble, M.R., Van Elst, L.T. (1999). On some clinical implications of the ventral striatum and the extended amygdala. Investigations of aggression. *Annals of the New York Academy of Sciences*, 877, 628–446. Review.

- Van Elst, L.T., Woermann, F.G., Lemieux, L., Thompson, P.J., Trimble, M.R. (2000). Affective aggression in patients with temporal lobe epilepsy: a quantitative MRI study of the amygdala. *Brain*, 123, 234–43.
- Veit, R., Flor, H., Erb, M., et al. (2002). Brain circuits involved in emotional learning in antisocial behavior and social phobia in humans. *Neuroscience Letters*, 328, 233–6.
- Volkow, N.D., Tancredi, L.R., Grant, C., et al. (1995). Brain glucose metabolism in violent psychiatric patients: A preliminary study. *Psychiatry Research: Neuroimaging*, 61, 243–53.
- Volkow, N.D., Tancredi, L.R. (1987). Neural substrates of violent behavior. A preliminary study with positron emission tomography. *British Journal of Psychiatry*, 151, 668–73.
- Vollm, B., Richardson, P., Stirling, J., et al. (2004). Neurobiological substrates of antisocial and borderline personality disorders: preliminary result of a functional fMRI study. *Criminal Behavior and Mental Health*, 14, 39–54.
- Woermann, F.G., Van Elst, L.T., Koepp, M.J., et al. (2000). Reduction of frontal neocortical grey matter associated with affective aggression in patients with temporal lobe epilepsy: an objective voxel by voxel analysis of automatically segmented MRI. *Journal of Neurology, Neurosurgery, and Psychiatry*, 68, 162–169.
- Wong, M.T., Fenwick, P.B., Lumsden, J., et al. (1997a). Positron emission tomography in male violent offenders with schizophrenia. *Psychiatry Research*, 68, 111–23.
- Wong, M.T., Lumsden, J., Fenton, G.W., Fenwick, P.B. (1997b). Neuroimaging in mentally abnormal offenders. *Issues Criminology and Legal Psychology*, 27, 49–58.
- Yang, Y., Raine, A., Lenz, T., Bihrl, S., LaCasse, L., Colletti, P. (2005). Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths. *Biological Psychiatry*, 57, 1103–8.
- Yang, Y., Raine, A., Narr, K.L., et al. (2006a). Successful and unsuccessful psychopaths: neuroanatomical similarities and differences. *Human Brain Mapping* [Abstract].
- Yang, Y., Raine, A., Narr, K.L., Lenz, T., Toga, A.W. (2006b). Amygdala volume reduction in psychopaths. *Society for Research in Psychopathology* [Abstract].

Ovulatory shifts in human female ornamentation: Near ovulation, women dress to impress

Martie G. Haselton^{a,*}, Mina Mortezaie^a, Elizabeth G. Pillsworth^b,
April Bleske-Rechek^c, David A. Frederick^d

^a Center for Behavior Evolution and Culture, Communication Studies and Department of Psychology, University of California, 2302 Rolfe Hall, Los Angeles, CA, 90095, USA

^b Center for Behavior Evolution and Culture and Department of Anthropology, University of California, Los Angeles, USA

^c Department of Psychology, University of Wisconsin-Eau Claire, USA

^d FPR-UCLA Center for Culture, Brain, and Development and Department of Psychology, University of California, Los Angeles, USA

Received 11 November 2005; revised 18 July 2006; accepted 20 July 2006

Available online 12 October 2006

Abstract

Humans differ from many other primates in the apparent absence of obvious advertisements of fertility within the ovulatory cycle. However, recent studies demonstrate increases in women's sexual motivation near ovulation, raising the question of whether human ovulation could be marked by observable changes in overt behavior. Using a sample of 30 partnered women photographed at high and low fertility cycle phases, we show that readily-observable behaviors – self-grooming and ornamentation through attractive choice of dress – increase during the fertile phase of the ovulatory cycle. At above-chance levels, 42 judges selected photographs of women in their fertile (59.5%) rather than luteal phase (40.5%) as “trying to look more attractive.” Moreover, the closer women were to ovulation when photographed in the fertile window, the more frequently their fertile photograph was chosen. Although an emerging literature indicates a variety of changes in women across the cycle, the ornamentation effect is striking in both its magnitude and its status as an overt behavioral difference that can be easily observed by others. It may help explain the previously documented finding that men's mate retention efforts increase as their partners approach ovulation.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Attractiveness; Clothing; Menstrual cycle; Ovulation; Sexual communication

One of the most noteworthy differences between humans and other closely related primates is the absence of clear advertisements of fertility within the ovulatory cycle (Dixson, 1998). Recent evidence has suggested, however, that there are subtle ovulatory cues in humans. Roberts et al. (2004) showed facial photographs of women taken during the follicular and luteal phases to male and female judges. On average, follicular phase images were judged more attractive approximately 54% of the time. Similarly, relative to those from other cycle phases, women's body scents near ovulation are judged as more attractive by men (Doty et al., 1975; Singh and Bronstad, 2001; Thornhill et al., 2003) and women's sexual desires vary across the cycle (Bullivant et al., 2004; Gangestad et al., 2002;

Haselton and Gangestad, 2006). Thus, human ovulation may not be completely concealed.

In the last decade, the literature on cyclic shifts in women's social motivations has grown rapidly. For ancestral women, the time required to collect food could have been considerable; thus, Fessler (2003) reasoned that there likely were tradeoffs across the cycle between feeding and other activities such as mating. Fessler (2003) compiled and reviewed evidence that women's appetites decrease near ovulation, and he hypothesized that this decrease in appetite at high fertility reflects an adaptation in women designed to decrease the motivational salience of goals that compete with efforts devoted to mating. As additional evidence supporting the hypothesis, Fessler reviewed studies showing that women's ranging activities, such as locomotion and volunteering for social activities, tend to increase near ovulation.

Other lines of evidence also indicate cyclic shifts in women's mating motivations. In a daily report study, Haselton and

* Corresponding author.

E-mail address: haselton@ucla.edu (M.G. Haselton).

Gangestad (2006) found that on high fertility days of the cycle women report a greater desire to go to clubs and parties where they might meet men. Macrae et al. (2002) found that women's ability to categorize male faces and male stereotypic words is faster near ovulation, suggesting increased attentiveness to "maleness" at high fertility. Other research shows that women's preferences for masculine features (e.g., masculine facial structure) increase near ovulation (reviewed in Gangestad et al., 2005a). Several rigorous within-subjects studies have found that women's attraction to and flirtation with men other than their primary partner is higher near ovulation than in other phases of the cycle (Gangestad et al., 2002; Haselton and Gangestad, 2006; also see Bullivant et al., 2004). Finally, Fisher (2004) found that women tested near midcycle, compared with those tested in other cycle phases, tend to give lower attractiveness ratings to photographs of female faces—an effect Fisher interpreted as evidence that women are more intrasexually competitive near ovulation. In sum, a variety of data sources indicate that women's social motivations – in particular, their sexual motivations – increase near ovulation.

Hypothesis: ovulation and ornamentation

We hypothesize that changes in women's motivations manifest themselves in changes in *self-ornamentation* through attentive personal grooming and attractive choice of dress. Ornamentation in non-humans, including bright plumage, long tails, and large bodies, is generally presumed to be the product of sexual selection (Andersson, 1994). These traits are effective in attracting mates, either because they indicate fitness (e.g., due to costs of their maintenance) or due to pre-existing sensory biases (Daly and Wilson, 1983; Parsons, 1995; Zahavi, 1975). Although rare, animals occasionally employ behavioral ornamentation, as opposed to morphological ornamentation, in the effort to attract mates. Male bowerbirds, for example, found in Australia and New Guinea, build elaborate structures ornamented with brightly colored flowers and fruits in order to attract mates. Male bowerbirds will often also pick up a brightly colored object in their beaks while displaying to a female, thus effectively ornamenting themselves (Diamond, 1982; Gilliard, 1969). The purpose of these traits, both morphological and behavioral, is to attract reproductive partners, and animals do not expend energy producing these displays when mating is not likely. The bowerbird dismantles its bower and abandons its territory during the non-breeding months (Pruett-Jones and Pruett-Jones, 1982), and even birds that rely on morphological ornamentation, such as brightly colored feathers or bills, may exhibit sexual dimorphism only seasonally (Badyaev and Duckworth, 2003; Peters et al., 2004).

In humans too, ornamentation may serve the purpose of attracting mates, at least in part. In a recent study, Grammer et al. (2005) interviewed women at a discotheque; those who rated their clothing as "sexy" and "bold" also reported that their intention for the evening was to flirt or find a sex partner. Although the direction of causality is unclear, these findings suggest that women's clothing choices are linked with their motivations.

Prior research also has shown that men's behaviors toward their partners shift across the cycle. Three studies have shown that, in the fertile relative to the luteal phase of the cycle, men are more attentive and loving toward their partners (Gangestad et al., 2002; Haselton and Gangestad, 2006; Pillsworth and Haselton, 2006) and two have shown that men are more jealous and possessive (Gangestad et al., 2002; Haselton and Gangestad, 2006). It is not yet known what cues drive these changes in men's behavior. One possibility is that men attend to differences in female behavior. For example, Haselton and Gangestad (2006) and Gangestad et al. (2002) found that women's reports of flirtatiousness with men other than their primary partner were higher when assessed during the late follicular as compared with the luteal phase of the cycle. In both studies, ovulatory increases in flirtatiousness statistically predicted ovulatory increases in male mate retention effort but did not fully account for them, leaving open the possibility that other ovulatory cues affect men's behavior—including the ornamentation effect we predict.

In this study, we measure an overt, readily observable behavior in women that we propose will be linked with ovulation. Specifically, we predict that women engage in self-ornamentation during the high fertility phase of the ovulatory cycle, thus placing themselves in the foreground of the social array.

Method

Procedure: photographic stimuli

Thirty women from the UCLA campus (mean age=21.07 years old; SD=2.35; range 18–37) posed for two standing full-body digital photographs with their hands placed at their sides (Canon PowerShot S410, 4.0 Megapixels). Women identified themselves as African American ($n=1$), Asian American ($n=10$), Caucasian ($n=6$), Hispanic/Latino ($n=7$), and mixed race or "other" ($n=6$). One photograph was taken on a high fertility day of the cycle (follicular phase) and one on a low fertility day of the cycle (luteal phase). Photographs were taken in the same location under standardized lighting conditions against a plain blue background. All women reported regular menstrual cycles (ranging between 26 and 35 days), were partnered (involved in a "committed romantic relationship" with a man), and none had used oral or other hormonal contraceptives within the last three months. Because previous studies have found stronger ovulatory effects in partnered than in non-partnered women (e.g., Havlicek et al., 2005; Pillsworth et al., 2004), we limited our investigation to partnered women.

Session scheduling and luteinizing hormone (LH) testing were conducted using the procedures described in Gangestad et al. (2002). There were three sessions—an initial session for cycle history assessment and scheduling and subsequent high and low fertility sessions. After initial sessions, women were scheduled to return for the next possible session (low or high) given their current cycle day. Low fertility sessions were scheduled to occur 4–10 days prior to the estimated day of next menstrual onset. Actual menstrual onset was reported by 66.7% of women after their low fertility session; for the balance of participants, menstrual onset was estimated using cycle length and the last date of menstrual onset. On average, based on these information sources, low fertility sessions took place 5.87 days prior to menses (SD=2.5; three women participated within 48 h of menstrual onset and possibly could have experienced premenstrual symptoms; therefore, days-to-menstrual-onset is included in the analyses presented below). High fertility sessions were scheduled to occur 15–17 days prior to the next estimated menstrual onset. Participants also reported to the laboratory to complete urine tests beginning two days prior to their high fertility session and continuing for three days after this session or until an LH surge was detected. Using an unmarked commercially available urinary stick ovulation test

(Clearblue™), all women were judged to have an LH surge between three days after and two days before their high fertility session. An LH surge typically proceeds ovulation by 24–48 h (Guermandi et al., 2001); thus, all women were likely to be near ovulation during their high fertility session. Within the fertile window of the cycle, conception risk increases as ovulation approaches (Wilcox et al., 1995). We therefore estimated days-to-ovulation (by adding two to days-to-LH surge; mean=3.03, SD=1.40) and included this estimate in the analyses reported below.

These 30 women were a subset of 58 originally recruited for the study. Women ineligible for inclusion in the study either showed no evidence of an LH surge ($n=4$), were rescheduled for low fertility sessions (due to their own time constraints) on days falling outside of the range of the luteal phase days ($n=3$), did not consent to having photos taken ($n=7$), consented to having their photos taken but did not consent to having their photos judged by people other than the researchers ($n=7$), or did not complete all sessions ($n=7$). There were no significant differences in relationship satisfaction, sociosexuality (Simpson and Gangestad, 1991), age, or relationship length between women retained in the study and those who were ineligible.

Participants were blind to the purpose of the study. They were told that the study examined health, personality, and sexuality, and that the urine tests examined “normal body chemistry.” They were told that the photographs were being taken to assess attractiveness and the accuracy of independent raters’ perceptions of their personality based on photographs alone; photographs were taken in each session ostensibly for reliability purposes. In the initial session, participants completed questionnaires that included information on sexuality and personality, thus justifying our description of the purposes of the study. In extensive debriefing, none guessed that the purpose of the study was to examine changes in clothing or attractiveness across the menstrual cycle.

To prevent any impact of variation in facial expression or facial appearance on ornamentation judgments, oval masks obscuring the entire face were applied to the photographs, leaving visible only hairstyle, jewelry, and clothing from head to toe.

Judges and experimental procedure

Judges were volunteers recruited by word of mouth (17 men and 25 women) from the UCLA and University of Wisconsin-Eau Claire campuses. Judges reported their ages by checking categories in a questionnaire: 28 were between 18 and 23 years old, seven were between 24 and 30, three were between 31 and 40, and two were 41 or older. All were blind to the fertility status of the photographs.

Image pairs were judged using a computer-based survey program (SurveyConsole™; <http://www.surveyconsole.com/>) that randomized the order of presentation of the 30 pairs and, within each pair, randomized the image (high vs. low fertility) presented on the left side of the screen. Judges were sent a unique Internet address link to the survey. Image sizes varied somewhat among judges depending on the size of the computer monitor used, with an approximate range of 4–6 in. high and 1.25–2 in. wide. For each set of images judges were asked, “In which photo is the person trying to look more attractive?” Judges mouse-clicked a box next to their choice. Judgments were made using this computer-based method to ensure that the fertility status of the photograph appearing on the right vs. the left side of the screen was random, thus ruling out the possibility that the lateral placement of high fertility images could not account for any effects we observed.

Results and discussion

We calculated scores for each woman by summing male and female judges’ choices of her high fertility image and converting it to a percentage. Thus, a woman whose high fertility photo was chosen by all judges would receive a score of 100%. There was high inter-judge agreement about whether each woman’s high or low fertility photograph was the one in which she was trying to look more attractive (agreement among male judges, $\alpha=0.86$; agreement among female judges, $\alpha=0.94$). We conducted a repeated measures analysis (General

Linear Models, SPSS 12.0) using photographed women as units of analysis (for information about judges as units of analysis, see below). To examine whether choices of high fertility images differed by sex of judge, sex of judge was a repeated factor (scores based on percent of male judges selecting the participant’s high fertility photo vs. scores based on percent of female judges selecting a participant’s high fertility photo). Days-to-ovulation (days from the date of a woman’s high fertility photograph to ovulation) and days-to-menstrual-onset (days from the date of a woman’s low fertility photograph to menstrual onset) were covariates.

This analysis includes a test of whether the overall marginal mean score (adjusted for covariates) differs from zero (a test of the intercept). We used this test to assess whether women’s overall scores differed from chance (50% high fertility choices). To do this, we unit-transformed women’s scores by subtracting 50% before conducting the analysis. The test of the intercept tells us whether the remainder is significantly above 0% (reflecting chance in the transformed scores). In the analysis, days-to-menstrual-onset was zero-centered, whereas days-to-ovulation was not transformed. Therefore, the effect of fertility status (testing whether a woman’s score was above chance) was estimated at the average of days prior to menstrual onset (roughly a mid-luteal phase day of the cycle) and at the day of ovulation. Scores were above chance, $F(1, 27)=7.06, p=0.013$ (mean=59.5%; SEM=5%).

There also was an effect of days-to-ovulation, $F(1, 27)=4.25, p=0.049$, such that, within the fertile window, women who were photographed closer to the day of ovulation had higher scores (see Fig. 1). There was no effect of days-to-menstrual onset on women’s scores, $F(1, 27)=0.01, \text{NS}$. Male and female judges’ scores did not differ, nor did sex of judge interact with days-to-ovulation ($F[1, 27]=0.66, \text{NS}$; $F[1, 27]=2.59, \text{NS}$, respectively). There was an interaction of sex of judge and days-to-menstrual-onset, $F(1, 27)=5.97, p=0.021$; although neither simple effect was significant ($p>0.50$), male judges tended to choose high fertility images more as women neared menses in their low fertility session, whereas the opposite was true for female judges.

In the facial attractiveness study noted in the introduction, Roberts et al. (2004) performed analyses treating judges as units of analysis rather than the photographed women as units of analysis as we have done. An analysis treating each judge as the unit of analysis would not permit generalization of an ovulatory cycle effect to the population from which our sample was drawn—the population of possible high vs. low fertility photographs—which is the primary population of interest. Rather, such an analysis examines whether results can be generalized to the population from which the judges were drawn. Nonetheless, analyses treating judges as units of analysis generated highly significant results.

The results support our prediction that women engage in active ornamentation, potentially in an attempt to appear more attractive, during the high fertility phase of the ovulatory cycle. Our inspection of the photographs suggested that this effect manifested itself in varied ways, some more obvious than others. For example, two women who dressed similarly in each session

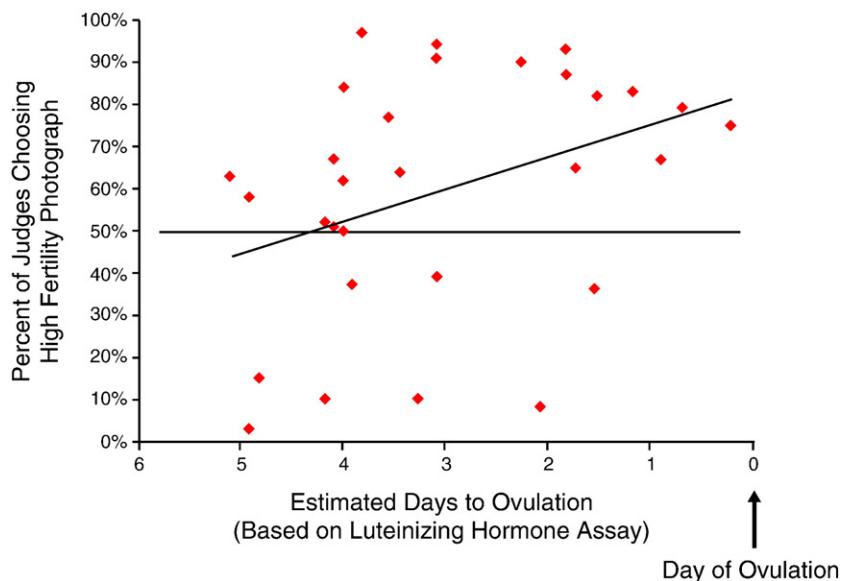


Fig. 1. Association between a woman's estimated number of days from ovulation when her high fertility photograph was taken and the percent of judges selecting it (controlling for the number of days between a woman's low fertility session and the date of next menstrual onset), partial $r(27)=0.37$, $p=0.049$.

wore tops with lace trim at high fertility; three wore skirts instead of pants; one woman added a fringy neck scarf; and several women simply showed more skin.

To systematically examine these ornamentation differences, three female research assistants coded the photographs by making eight qualitative assessments related to ornamentation (see Table 1). All coders were blind to the fertility status of the photographs. Two made initial assessments and the third resolved discrepancies. Coders were asked to indicate, for example, the photo in which the woman wore “more fashionable clothing” or “showed more skin” (possible responses: right photo, left photo, or no difference). Given low cell counts for some high versus low fertility comparisons (e.g., wearing a skirt in one session but not the other, high fertility = 3 women; low fertility = 0 women), and given that we had already conducted statistical comparisons of the high and low fertility photographs, we did not conduct statistical tests on

these secondary codings. The results, however, are clear. High fertility clothing choices were coded as “more fashionable,” “nicer,” and showing more skin. Although women revealed more skin, clothing choices were not coded as “sexier” at high fertility, which could reflect norms of daytime student dress, in which sexually explicit clothing is not appropriate. Alternatively, the effect we have documented might reveal a general desire to look more attractive rather than to appear “sexy.”

As shown in Fig. 1, women who were photographed closer to ovulation, as indicated by the LH assay, had their high fertility photo chosen more often by judges. This demonstrates that choices of fertile images are due to differences attributable to changes in body chemistry related to ovulation and not due to differences attributable to proximity to menses (where, although there was substantial variation in proximity to menses in the low fertility session, there was no hint of an association of days-to-menstrual onset with women's scores). Also as shown in Fig. 1, four fertile phase photos are notable because more than 90% of judges did *not* choose them. These data demonstrate that dress is influenced, of course, by more than cycle phase. A variety of external constraints (e.g., whether a woman rushed to campus for an exam or had a job interview on the day in question) affect clothing choices as well. Given these constraints, the relatively large difference between judges' choices of high and low fertility images (59.5% vs. 41.5%) is impressive.

Our results support a similar prediction made by Grammer et al. (2005) that women at a discotheque would display more skin when in the fertile phase of the cycle. Grammer et al. (2005) found that women's stated sexual motivation was linked to the revealingness of the clothing they wore to the discotheque but did not find compelling evidence of an ovulatory effect. In some analyses, Grammer et al. found evidence that estradiol was correlated with clothing revealingness. However, because estradiol varies more between women than within women (Gann et al., 2001; Ferrell et al., 2005), and shows a secondary

Table 1
Differences in women's clothing choices across high and low fertility photographs

Judgment (percent concordance of judges' codes)	High fertility	Low fertility
Wearing “more fashionable clothes” (70%)	18	8
Wearing “nicer clothes” (77%)	17	8
Showing more skin (upper body) (77%)	11	6
Showing more skin (lower body) (93%)	7	5
Wearing “sexier clothes” (70%)	6	7
Wearing more “accessories” (63%)	6	7
Wearing a skirt in one session but not other (100%)	3	0
Wearing a lacy top (80%)	3	1

Values represent counts within each category. High and low fertility values do not sum to 30 because judges did not perceive a difference between all photographs on each dimension (e.g., 4 out of 30 photo pairs were not judged to differ on wearing “more fashionable clothes”). Percent concordance is the agreement between two female judges before the third female judge resolved discrepancies.

peak after ovulation (e.g., Allende, 2002), estradiol levels cannot be used to assess cycle phase in cross-sectional studies (cf. DeBruine et al., 2005). Grammer's results may reflect trait-level estrogen effects, as estrogen is associated with factors relevant to female mating strategies, such as attractiveness (Fink and Penton-Voak, 2002; Law Smith et al., 2006) and age (Ferrell et al., 2005).

There are several competing explanations for the link between ovulation and clothing choices that we have documented. First, emerging evidence indicates that women become subtly more attractive near ovulation (e.g., Roberts et al., 2004; Singh and Bronstad, 2001; Thornhill et al., 2003). It is possible that women are sensitive to cyclic changes in their own physical attractiveness (Haselton and Gangestad, 2006), and they may choose to display their attributes through ornamentation that can result in any of a variety of attractiveness-related benefits, including esteem in the eyes of peers and increased attention from men (Langlois et al., 2000).

Second, as outlined in the introduction, women's mating motivations vary across the cycle (e.g., Bullivant et al., 2004; Fessler, 2003; Gangestad et al., 2002; Haselton and Gangestad, 2006). Thus, women may engage in ornamentation to attract mates. It is possible, in particular, that women may be aiming to attract mates *other than* their primary partners in order to gain access to good genes. Several rigorous within-subjects studies have found increases in women's desires for men other than their primary partner at high fertility (Gangestad et al., 2002, 2005b; Haselton and Gangestad, 2006; Pillsworth and Haselton, 2006), but no changes in desires for the primary partner (Gangestad et al., 2002, 2005b; Pillsworth and Haselton, 2006). Three studies show that increases in extra-pair attractions at high fertility are greatest for women who could, in theory, gain the most from extra-pair mating for genetic benefits, specifically women whose primary mates are relatively asymmetrical (Gangestad et al., 2005b) or lacking sexual attractiveness (Haselton and Gangestad, 2006; Pillsworth and Haselton, 2006). Preliminary evidence also suggests that women are intrasexually competitive near ovulation (Fisher, 2004), possibly because this is the time at which the costs and benefits of mating decisions are greatest (Fessler, 2003). Thus, near ovulation, women may also be competing with other women for the best mating opportunities.

In addition to these possibilities, changes in style of dress could reflect more general changes across the cycle, such as changes in mood. The evidence of cyclic changes in mood is mixed. Some studies show changes that are limited to or most prominent in women with premenstrual syndrome (Metcalf and Livesey, 1995; Sanders et al., 1983; also see Bäckström, 1983). Other studies fail to find any significant changes in mood across the cycle (e.g., Laessle et al., 1990; Van Goozen et al., 1997; Lahmeyer et al., 1982; Patkai et al., 1974). Several studies have shown that changes in mood across the cycle may be attributable to women's expectancies (leading to biased self-reporting) rather than genuine effects of the cycle on mood (e.g., Olasov and Jackson, 1987; Weidner and Helmig, 1990; also see Ruble, 1977). To the extent that mood effects are real, most evidence points to increases in negative affect as menstrual onset

approaches (e.g., Bäckström, 1983; Dennerstein and Burrows, 1979; Sanders et al., 1983; Van Goozen et al., 1997). If changes in mood are responsible for differences in women's style of dress from high to low fertility, one might naturally expect an increase in judges' choices of high fertility images for women whose low fertility photographs were taken closer to the day of menstrual onset, which we examined but did not find. Rather, the effects appear to be driven by proximity to ovulation.

Women also may become more sociable in general near ovulation. For example, although evidence is somewhat mixed (see Fessler, 2003) several studies suggest that women engage in more locomotor behavior near ovulation (measured using pedometers). If ovulation leads women to feel more sociable, they may also attend more to personal grooming and style of dress.

The ornamentation effect could be explained by some combination of the factors we have described. The mood and sociability effects, for example, could be the proximal psychological motives that underpin mating functions of attempting to appear more attractive. Indeed, Fessler (2003) proposed that evidence of periovulatory decreases in feeding behavior reflect adaptations for reducing the motivational salience of activities that compete directly or indirectly with mating, and periovulatory increases in social behavior reflect increased effort allocated to mating activities.

Each of the causal pathways we have outlined above can be tested in future research. For example, the proposal that ornamentation differences are directed toward extra-pair mates could be tested by examining whether partnered women show the effect more than unpartnered women and by examining whether the effect is attenuated among women whose long-term mates are attractive and symmetrical. The intrasexual competition function could be tested by subtly leading women to anticipate that their laboratory sessions will involve other female participants (intrasexual competitors) and will be run by highly attractive male experimenters.

Regardless of what underlying motivations or evolved functions are tied to the ornamentation effect, our results appear to provide strong objective evidence of changes in women's overt, observable behaviors associated with ovulation. In contrast to subtle changes in facial appearance (Roberts et al., 2004) and body scents (Doty et al., 1975; Singh and Bronstad, 2001; Thornhill et al., 2003), variation in women's self-grooming and ornamentation behaviors are perhaps the most readily available cues to ovulation and associated shifts in female motivation available to male partners. They may therefore be responsible, in part, for changes in relationship dynamics across the cycle, such as ovulatory increases in men's mate retention efforts (Gangestad et al., 2002; Haselton and Gangestad, 2006; Pillsworth and Haselton, 2006). The ornamentation effect is one of the most striking pieces of evidence that ovulation in humans is not fully concealed.

Acknowledgments

This research was supported by a UCLA Council on Research Faculty Grant. We are grateful to Kristina Grigorian, Lisa Newon, Carolyn Pratt, Jean Torres, and Andrea Yocom for

assistance in collecting the data and to Unipath Diagnostics for generously donating LH tests. We thank Steve Gangestad for his expert statistical advice.

References

- Allende, M.E., 2002. Mean versus individual hormonal profiles in the menstrual cycle. *Fertil. Steril.* 78, 90–95.
- Andersson, M.B., 1994. Sexual Selection. Princeton Univ. Press, Princeton.
- Bäckström, T., 1983. Mood, sexuality, hormones, and the menstrual cycle: II. hormone levels and their relationship to the premenstrual syndrome. *Psychosom. Med.* 45, 503–507.
- Badyaev, A.V., Duckworth, R.A., 2003. Context-dependent sexual advertisement: plasticity in development of sexual ornamentation throughout the lifetime of a passerine bird. *J. Evol. Biol.* 16, 1065–1076.
- Bullivant, S.B., Sellergren, S.A., Stern, K., Spencer, N.A., Jacob, S., Mennella, J.A., McClintock, M.K., 2004. Women's sexual experience during the menstrual cycle: identification of the sexual phase by noninvasive measurement of luteinizing hormone. *J. Sex Res.* 41, 82–93.
- Daly, M., Wilson, M., 1983. Sex, Evolution, and Behavior, 2nd ed. Willard Grant Press, Boston.
- DeBruine, L.M., Jones, B.C., Perrett, D.I., 2005. Women's attractiveness judgments of self-resembling faces change across the menstrual cycle. *Horm. Behav.* 47, 379–383.
- Dennerstein, L., Burrows, G.D., 1979. Affect and the menstrual cycle. *J. Affective Disord.* 1, 77–92.
- Diamond, J.M., 1982. Evolution of bowerbirds' bowers: animal origins of the aesthetic sense. *Nature* 297, 99–102.
- Dixson, A., 1998. Primate Sexuality: Comparative Studies of the Prosimians, Monkeys, Apes, and Human Beings. Oxford Univ. Press, New York.
- Doty, R.L., Ford, M., Preti, G., Huggins, G.R., 1975. Changes in the intensity and pleasantness of human vaginal odors during the menstrual cycle. *Science* 190, 1316–1317.
- Ferrell, R.J., O'Connor, K.A., Rodriguez, G., Gorriño, T., Holman, D.J., Brindle, E., Miller, R.C., Schechter, D.E., Korshalla, L., Simon, J.A., Mansfield, P.K., Wood, J.W., Weinstein, M., 2005. Monitoring reproductive aging in a 5-year prospective study: aggregate and individual changes in steroid hormones and menstrual cycle lengths with age. *Menopause* 12, 567–577.
- Fessler, D.M.T., 2003. No time to eat: an adaptationist account of periovulatory behavioural changes. *Q. Rev. Biol.* 78, 3–21.
- Fink, B., Penton-Voak, I.S., 2002. Evolutionary psychology of facial attractiveness. *Curr. Dir. Psychol. Sci.* 11, 154–158.
- Fisher, M.L., 2004. Female intrasexual competition decreases female facial attractiveness. *Proc. Biol. Sci. (Suppl.)* 271, 283–285.
- Gangestad, S.W., Thornhill, R., Garver, C., 2002. Changes in women's sexual interests and their partners' mate retention tactics across the menstrual cycle: evidence for shifting conflicts of interest. *Proc. R. Soc. London, B* 269, 975–982.
- Gangestad, S.W., Thornhill, R., Garver-Apgar, C.E., 2005a. Adaptations to ovulation. In: Buss, D.M. (Ed.), The Handbook of Evolutionary Psychology. John Wiley and Sons, Hoboken, NJ, pp. 344–371.
- Gangestad, S.W., Thornhill, R., Garver-Apgar, C.E., 2005b. Women's sexual interests across the ovulatory cycle depend on primary partner fluctuating asymmetry. *Proc. R. Soc. London, B* 272, 2023–2027.
- Gann, P.H., Giovanazzi, S., Van Horn, L., Branning, A., Chatterton, R.T., 2001. Saliva as a medium for investigating intra- and interindividual differences in sex hormone levels in premenopausal women. *Cancer Epidemiol., Biomarkers Prev.* 1, 59–64.
- Grammer, K., Renninger, L., Fischer, B., 2005. Disco clothing, female sexual motivation, and relationship status: is she dressed to impress? *J. Sex Res.* 41, 66–74.
- Giilliard, E.T., 1969. Birds of Paradise and Bower Birds. Natural History Press, Garden City.
- Guermanni, E., Veggetti, W., Bianchi, M.M., Ugliesti, A., Ragni, G., Crosignani, P., 2001. Reliability of ovulation tests in infertile women. *Obstet. Gynecol.* 97, 92–96.
- Haselton, M.G., Gangestad, S.W., 2006. Conditional expression of women's desires and men's mate guarding across the ovulatory cycle. *Horm. Behav.* 49, 509–518.
- Havlicek, J., Roberts, S.C., Flegl, J., 2005. Women's preference for dominant male odour: effects of menstrual cycle and relationship status. *Biol. Lett.* 1, 256–259.
- Laessle, R.G., Tuschl, R.J., Schweiger, U., Pirke, K.M., 1990. Mood changes and physical complaints during the normal menstrual cycle in healthy young women. *Psychoneuroendocrinology* 15, 131–138.
- Lahmeyer, H.W., Miller, M., DeLeon-Jones, F., 1982. Anxiety and mood fluctuation during the normal menstrual cycle. *Psychosom. Med.* 44, 183–194.
- Langlois, J.H., Kalakanis, L., Rubenstein, A.J., Larson, A., Hallam, M., Smoot, M., 2000. Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol. Bull.* 126, 390–423.
- Law Smith, M.J., Perrett, D.I., Jones, B.C., Cornwell, R.E., Moore, F.R., Feinberg, D.R., Boothroyd, L.G., Stirrat, M.R., Whiten, S., Pitman, R.M., Hillier, S.G., 2006. Facial appearance is a cue to reproductive hormone levels in women. *Proc. R. Soc. London, B* 273 (1583), 135–140.
- Macrae, C.N., Alnwick, K.A., Milne, A.B., Schloerscheidt, A.M., 2002. Person perception across the menstrual cycle: hormonal influences on social-cognitive functioning. *Psychol. Sci.* 13, 532–536.
- Metcalf, M.G., Livesey, J.H., 1995. Distribution of positive moods in women with the premenstrual syndrome and in normal women. *J. Psychosom. Res.* 39, 609–618.
- Olasov, B., Jackson, J., 1987. Effects of expectancies on women's reports of moods during the menstrual cycle. *Psychosomatic Med.* 49, 65–78.
- Parsons, P.A., 1995. Stress and limits to adaptation: sexual ornaments. *J. Evol. Biol.* 8, 455–461.
- Patkai, P., Johannsson, G., Post, B., 1974. Mood, alertness and sympathetic-adrenal medullary activity during the menstrual cycle. *Psychosomatic Med.* 36, 503–512.
- Peters, A., Denk, A.G., Delhey, K., Kempenaers, B., 2004. Carotenoid-based bill colour as an indicator of immunocompetence and sperm performance in male mallards. *J. Evol. Biol.* 17, 1111–1120.
- Pillsbury, E.G., Haselton, M.G., 2006. Male sexual attractiveness predicts differential ovulatory shifts in female extra-pair attraction and male mate retention. *Evol. Hum. Behav.* 27, 247–258.
- Pillsbury, E.G., Haselton, M.G., Buss, D.M., 2004. Ovulatory shifts in female sexual desire. *J. Sex Res.* 41, 55–65.
- Pruett-Jones, M.A., Pruett-Jones, S.G., 1982. Spacing and distribution of bowers in Macgregor's bowerbird (*Amblyornis macgregoriae*). *Behav. Ecol. Sociobiol.* 11, 25–32.
- Roberts, S.C., Havlicek, J., Flegl, J., Hruskova, M., Little, A.C., Jones, B.C., Perrett, D.I., Petrie, M., 2004. Female facial attractiveness increases during the fertile phase of the menstrual cycle. *Proc. R. Soc. London, B (Suppl.)* 271, S270–S272.
- Ruble, D.N., 1977. Premenstrual symptoms: a reinterpretation. *Science* 197, 291–292.
- Sanders, D., Warner, P., Bäckström, T., Bancroft, J., 1983. Mood, sexuality, hormones and the menstrual cycle: I. changes in mood and physical state: description of subjects and method. *Psychosom. Med.* 45, 487–501.
- Simpson, J.A., Gangestad, S.W., 1991. Individual differences in sociosexuality: evidence for convergent and discriminant validity. *J. Pers. Soc. Psy.* 60, 870–883.
- Singh, D., Bronstad, P.M., 2001. Female body odour is a potential cue to ovulation. *Proc. R. Soc. London, B* 268, 797–801.
- Thornhill, R., Gangestad, S.W., Miller, R., Scheyd, G., McCollough, J.K., Franklin, M., 2003. Major histocompatibility complex genes, symmetry, and body scent attractiveness in men and women. *Behav. Ecol.* 14, 668–678.
- Van Goozen, S.H.M., Wiegant, V.M., Endert, E., Helmond, F.A., 1997. Psychoendocrinological assessment of the menstrual cycle: the relationship between hormones, sexuality, and mood. *Arch. Sex. Behav.* 26, 359–382.
- Weidner, G., Helming, L., 1990. Cardiovascular stress reactivity and mood during the menstrual cycle. *Women Health* 16, 5–21.
- Wilcox, A.J., Weinberg, C.R., Baird, D.D., 1995. Timing of sexual intercourse in relation to ovulation: effects of probability of conception, survival of pregnancy, and sex of the baby. *N. Engl. J. Med.* 333, 1517–1521.
- Zahavi, A., 1975. Mate selection: a selection for a handicap. *J. Theor. Biol.* 53, 205–214.

A neural basis for the effect of candidate appearance on election outcomes

Michael L. Spezio,^{1,2,*} Antonio Rangel,^{2,*} Ramon Michael Alvarez,² John P. O'Doherty,^{2,3} Kyle Mattes,^{2,4} Alexander Todorov,⁵ Hackjin Kim,^{2,6} and Ralph Adolphs²

¹Scripps College, 1030 Columbia Avenue, Claremont, CA 91711, ²California Institute of Technology, M/C 228-77, Pasadena, CA 91101,

³Trinity College Institute of Neuroscience, Trinity College, Dublin, Ireland, ⁴Department of Political Science, University of Iowa, Iowa City, IA 52240, ⁵Department of Psychology, Princeton University, Princeton, NJ 08540, USA and ⁶Korea University, Psychology, Seoul, Korea 136-701

Election outcomes correlate with judgments based on a candidate's visual appearance, suggesting that the attributions viewers make based on appearance, so-called thin-slice judgments, influence voting. Yet, it is not known whether the effect of appearance on voting is more strongly influenced by positive or negative attributions, nor which neural mechanisms subserve this effect. We conducted two independent brain imaging studies to address this question. In Study 1, images of losing candidates elicited greater activation in the insula and ventral anterior cingulate than images of winning candidates. Winning candidates elicited no differential activation at all. This suggests that negative attributions from appearance exert greater influence on voting than do positive. We further tested this hypothesis in Study 2 by asking a separate group of participants to judge which unfamiliar candidate in a pair looked more attractive, competent, deceitful and threatening. When negative attribution processing was enhanced (specifically, under judgment of threat), images of losing candidates again elicited greater activation in the insula and ventral anterior cingulate. Together, these findings support the view that negative attributions play a critical role in mediating the effects of appearance on voter decisions, an effect that may be of special importance when other information is absent.

Keywords: social cognition; faces; trait judgment; voting; political science

We rapidly evaluate others based on their appearance, an effect that has been well demonstrated in social psychology (Hassin and Trope, 2000; Todorov and Uleman, 2003; Willis and Todorov, 2006) and not lost on political scientists (Rosenberg, 1986; Rosenberg *et al.*, 1991). In fact, recent behavioral studies have shown that judgments about candidates' physical appearance correlate with real election outcomes (Todorov *et al.*, 2005; Ballew and Todorov, 2007), suggesting that information derived from visual appearance alone (so-called thin-slice information) affects voting behavior. In particular, work done by Todorov and coworkers showed that competence judgments made about candidates who ran in 2006 gubernatorial elections picked out the winner at significantly above-chance levels ($57 \pm 6\%$, mean \pm s.d.) (Ballew and Todorov, 2007). This replicated an earlier study, where the average association between an individual viewer's competence judgments and electoral victory ranged from $53 \pm 10\%$ to $59 \pm 7\%$ for various types

of political office (Todorov *et al.*, 2005). Interestingly, participants made these judgments about politicians with whom they were unfamiliar, and after only very brief exposures to the images (100–2000 ms). These findings are quite remarkable given the amount of information about candidates to which a typical voter is exposed. For example, in the 2006 U.S. midterm elections, candidates and their interest groups spent over a billion dollars on advertising to inform voters of their party affiliation, record, policies and personal qualities (CNN, 2006). While some political scientists view such rich information as the primary driver of voter decisions (Popkin, 1991; Prior, 2005), there are other data suggesting that voters make use of much sparser information (Downs, 1957; Alvarez, 1997).

Given that mere appearance seems to influence voting behavior, this raises the question of what psychological and neural processes might mediate this effect. Here we investigate whether the effect of appearance on voting is more strongly influenced by positive or negative attributions (which might be either implicit or explicit; Galdi *et al.*, 2008). Given the association between appearance-based judgments about competence (a putatively positive trait) and electoral victory (Todorov *et al.*, 2005), one might expect that positive attributions from a politician's appearance dominate in influencing decisions by voters. Yet there are several reasons to hypothesize that negative attributions

Received 6 October 2008; Accepted 17 October 2008

Advance Access publication 28 October 2008

The authors wish to thank Jan Glaescher for helpful discussions and for the use of *rxfplot*. This research was financed in part by a grant from the Gordon and Betty Moore Foundation, and with support from the National Science Foundation (grant number SES-0134618 to A.R.) and from the National Institutes of Health (to R.A. and M.L.S.).

*These authors contributed equally to this work.

The authors declare that they have no competing financial interests.

Correspondence should be addressed to Dr Michael L. Spezio, Scripps College, 120 Steele Hall, 1030 Columbia Avenue, Claremont, CA 91711, USA. E-mail: mspezio@scrippscollege.edu.

also play a role. First, one would predict that thin-slice attributions play a larger role in influencing voters who have little knowledge of candidates, and there is strong evidence that, when voters know little about a candidate, perceived negative aspects of a candidate exert a stronger influence than positive aspects on voter turnout (Lau and Pomper, 2001; Martin, 2004; Stevens *et al.*, 2008), party defection, the number of self-reported reasons for voting for or against, and predicting overall positive and negative evaluations of a candidate (for a review, see Lau, 1982). Second, voters make implicit ingroup/outgroup distinctions about candidates (Iyengar and Simon, 2000), and negative information dominates the evaluation of outgroup members (Forgas and Fiedler, 1996). To the extent that candidates occupy a default outgroup position for voters lacking relevant information, as suggested by evidence showing that increasing a candidate's facial similarity to a voter significantly increases that voter's support (Bailenson *et al.*, 2006), negative attributions will dominate voter decisions. Third, negative motives play a significant role in voter decisions, particularly in so-called negative voting, in which a vote for Candidate A is really just a vote against Candidate B (Kernell, 1977; Fiorina and Shepsle, 1989).

There is, thus, reason to hypothesize that candidate appearance affects voting through processes that evaluate both positive as well as negative traits, and that may be both implicit and explicit. To test whether positive or negative attributions play a primary role, and to identify the neural regions involved, we conducted two fMRI studies using images of real political candidates. The first study

asked participants to vote in the scanner and the second asked an independent group of participants to make both positive and negative trait judgments. In the first study, we would expect the particular brain structures activated during voting to provide insight into the relative influences of the above factors. Specifically, if voting is more influenced by positive attributions, we would expect vote winners to elicit greater activation of brain structures known to be involved in affective processing. Conversely, if negative attributions dominate, vote losers should elicit the greater activation in these structures. Following up on the results from this first study, our second study used explicit trait judgments (threat, deceitfulness, attractiveness, competence) to further investigate the relative contributions made by positive and negative attributions in linking candidate appearance to voter decisions. We expected that similar neural mechanisms would be revealed, further supporting the findings from the first study.

RESULTS

Study 1: simulated voting

In the first study, participants voted in a simulated election. Participants viewed grayscale pictures of 100 pairs of unfamiliar real politicians, one Republican and one Democrat, who competed in the 2006 U.S. midterm elections (Figure 1A; see Methods for details). Participants saw the image of each of the two candidates in sequence, for only 1 s, separated by a blank screen, and after a delay were asked to cast their vote. All the data were collected before the 2006 election, ensuring that participants could not have been influenced in any way by the real election outcomes.

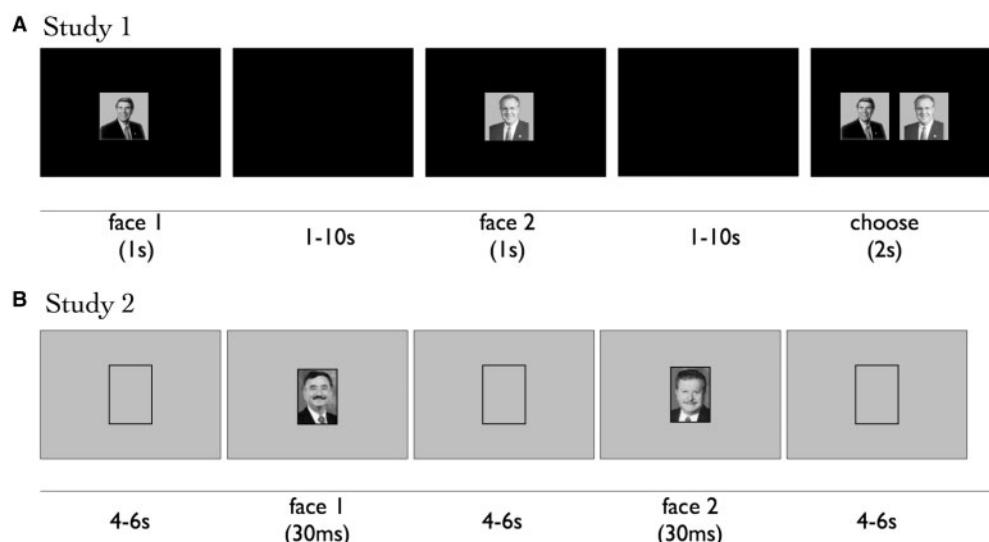


Fig. 1 Trial design of the two studies. Each experiment showed participants grayscale images of real politicians, one at a time, separated by a variable interstimulus interval. For both experiments, stimuli are depicted as they appeared on the screen. (A) Study 1: simulated voting study. Participants were shown grayscale images of each of the candidates, separately, for 1 s each. The images were separated from each other and from the decision period by a 1–10 s blank screen. (B) Study 2: social judgment study. Each trial consisted of two cycles of the alternating presentation of two images, separated by blank boxes that cued viewers to the subsequent location of the image. Participants were asked to indicate which of the two images looked more threatening, attractive, deceitful or competent (in four separate blocks) by pushing one of two buttons (whose location was indicated by a small black dot in the lower corner of the screen). Images were shown for 30 ms (unmasked). We analyzed data only from the first cycle (shown in the figure), but not from the second cycle, to ensure consistency with our first study and to maximize association with thin-slice processing.

In order to increase the sensitivity of our fMRI analyses, we limited our analyses to voxels within brain areas already known to be associated with the evaluation of facial appearance and affective processing. We produced region of interest (ROI) masks using the Automated Anatomical Labeling Toolbox for SPM (Tzourio-Mazoyer *et al.*, 2002). These ROIs are as follows: (i) the bilateral temporal lobes, including the fusiform gyrus (associated with gaze and face processing; Kanwisher and Yovel, 2006); (ii) the bilateral caudate (associated with positive evaluation of faces; Kim *et al.*, 2007); (iii) the bilateral putamen (associated with reward-based processing; O'Doherty *et al.*, 2004); (iv) the bilateral gyrus rectus (associated with positive evaluation of faces; Kim *et al.*, 2007); (v) the bilateral orbitofrontal cortex (associated with positive evaluation of faces; Gottfried *et al.*, 2003; Kim *et al.*, 2007); (vi) the bilateral insula [associated with perception of lack of trustworthiness (Winston *et al.*, 2002), and pain perception (Ploghaus *et al.*, 1999; Salomons *et al.*, 2007), but also with positive evaluation of faces (Kim *et al.*, 2007)]; (vii) the bilateral amygdala (associated mostly with negative facial attribution; Winston *et al.*, 2002); and (viii) the bilateral anterior cingulate (associated with social rejection; Eisenberger *et al.*, 2003; Somerville *et al.*, 2006). We report only those clusters surviving FWE correction at $P < 0.05$, as determined by Monte Carlo simulation using AlphaSim in AFNI (Xiong *et al.*, 1995; Cox, 1996) (see Supporting Information).

We analyzed the fMRI data for this first study by using the voting data provided by the participants in our experiment. We estimated a general linear model in which the appearance of an individual candidate's image was modulated by its vote share in the simulated election, a variable that we refer to as lab vote share (see Methods for details). We found that positive lab vote share elicited no significant activation in any ROIs. In fact, positive lab vote share did not result in any significant activation anywhere, even with a whole-brain analysis. In contrast, negative lab vote share (i.e. election loss) elicited robust, statistically significant activations in bilateral insula [222 voxels (48, -3, -9), $Z = 3.81$; 179 voxels (-45, 12, 9); $Z = 3.96$; Supporting Information Table S1; Figure 2A], and bilateral anterior cingulate cortex [239 voxels (3, 33, 9), $Z = 3.73$; Supporting Information Table S1; Figure 3A]. Thus, these regions are increasingly engaged by viewing candidates with larger margins of electoral loss.

The lack of any significant brain activations elicited by viewing winners in our simulated vote, coupled with the robust activations elicited by viewing losers, suggests that negative attributions from appearance may play a predominant role in mediating how appearance influences voting.

Study 2: candidate trait judgments

In our second study, we sought to further investigate the relative contributions to this effect made by positive and negative attributions (which might be either implicit or

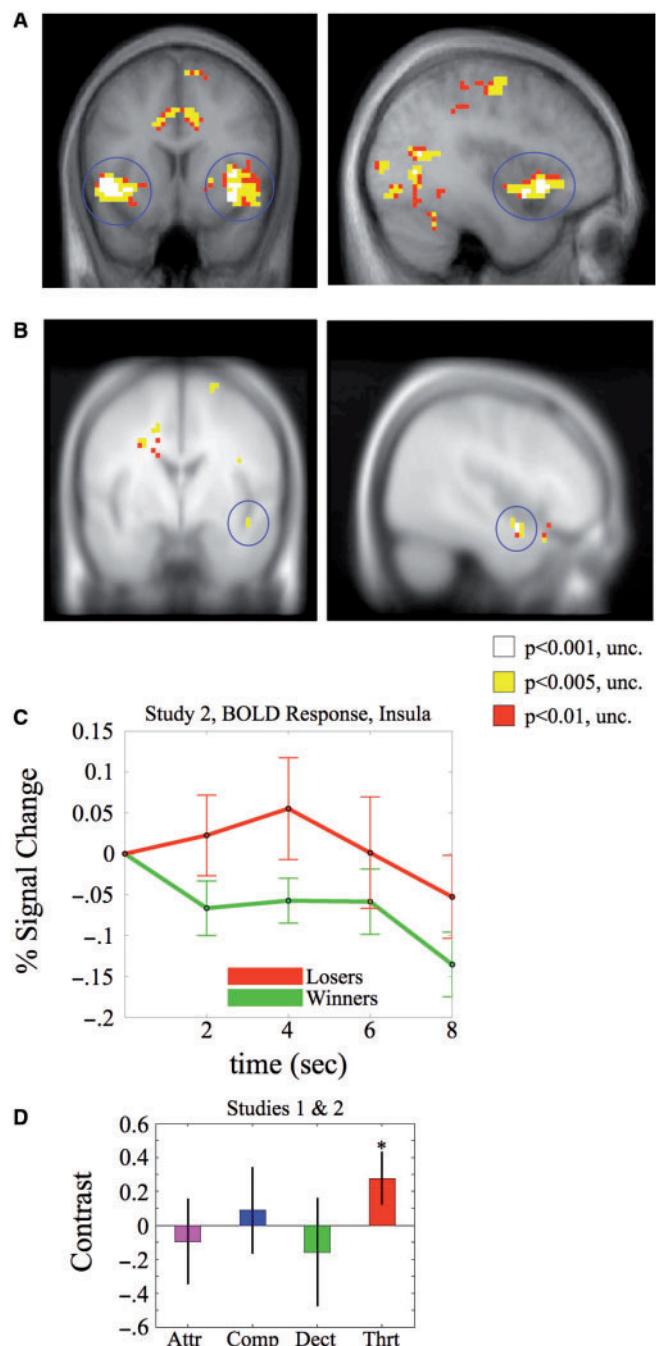


Fig. 2 Activation in the insula/parainsula correlates with election loss in both studies. (A) Study 1: activation during the simulated voting study in bilateral insula [blue circles; left insula (-45, 12, 9) and right insula (48, -3, -9)] was negatively correlated with lab vote share (greater for losers in the simulated election). (B) Study 2: activations elicited by images of candidates who lost real elections, for the contrast of loser > winner, under the condition of threat judgment. Shown are activations in the right insula, circled in blue. (C) Study 2: time course of activations for the peak voxel in the area of activation in the right insula, shown for real-electoral losers (red) and winners (green). Candidates who lost elicited an increase in activation in the right insula, while those who won actually elicited a decrease, consistent with our interpretation that this activation reflects negative attributions (means and s.e.m.). (D) Studies 1 and 2: group mean contrast estimates for loser > winner in real elections, under the four judgment conditions in Study 2, within a region in the right insula defined by the peak contrast in Study 1. Only the threat judgment (Thrt; red bar) shows significant effects (means and s.e.m.; * indicates $P < 0.05$).

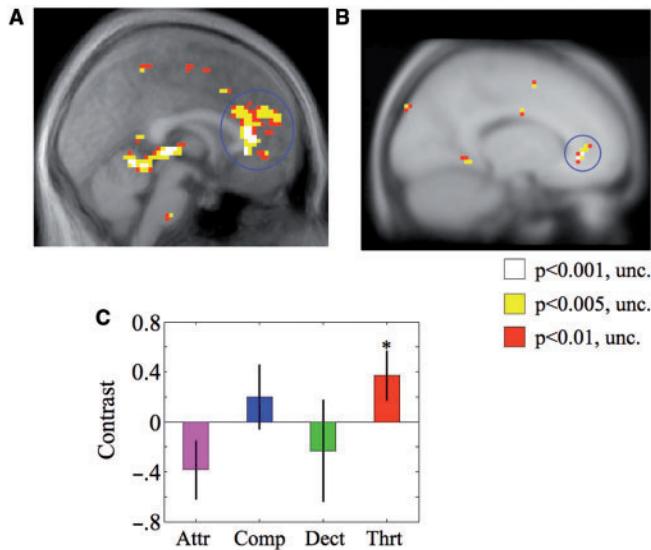


Fig. 3 Activation in ventral anterior cingulate cortex correlates with election loss in both studies. (A) Study 1: activation during the simulated voting study in bilateral anterior cingulate [blue circle; (3, 33, 9) and (-9, 21, 24)] was negatively correlated with lab vote share. (B) Study 2: activation elicited by images of candidates who lost real elections, for the contrast loser > winner (circled in blue), under the threat judgment condition. Activations are seen in the right ventral anterior cingulate. (C) Studies 1 and 2: group mean contrast estimates for loser > winner in real elections for the four judgment conditions in Study 2, within a region in the right anterior cingulate defined by the peak contrast in Study 1. Only the threat judgment (Thrt; red bar) shows significant effects (means and s.e.m. * indicates $P < 0.05$).

explicit), as well as the underlying neural structures we had found in the first study. In order to selectively enhance processing of one attribution over another, we asked participants to make overt judgments in the scanner. To best investigate the effects of candidate appearance alone, we maximized thin-slice conditions with a rapid stimulus presentation. We used a subset of the same images previously shown to elicit an association between real-electoral outcome and judgments of competence (Todorov *et al.*, 2005).

Study participants made social judgments based on the images of real, but unfamiliar, political candidates who ran against one another in the 2000, 2002 and 2004 U.S. Congressional elections, mainly in the House of Representatives. Participants made binary judgments about 30 pairs of candidates, one Republican and one Democrat (in randomized order), on two putatively positive traits, attractiveness (Attr) and competence (Comp), and two putatively negative traits, public deceitfulness (Dect) and personal threat (Thrt), in four separate scanning sessions (see Methods). Each trial consisted of a protocol that has been previously used to investigate face preferences (Kim *et al.*, 2007), in which the two images in a pair of candidates were presented sequentially, for only 30 ms each (unmasked), one alternating with the other, until the participant pushed a button to indicate which of the two faces showed more of the trait being judged (Figure 1B).

As one would expect if participants made meaningful judgments, positive trait judgments were positively

correlated (Attr and Comp, $r=0.39$, $P=0.002$), those for negative traits were positively correlated (Thrt and Dect, $r=0.61$, $P<0.0001$) and those between Comp and Thrt were negatively correlated (Comp and Thrt, $r=-0.39$, $P=0.002$). No other statistically reliable relationships were seen in the behavioral data alone.

The relationship between our behavioral data regarding competence judgments and real-world electoral outcome was in line with the published findings we reviewed above (Todorov *et al.*, 2005; Ballew and Todorov, 2007): we found that our participants were above chance in judging winners of real elections as more competent [$55 \pm 9\%$, $t(15)=2.15$, $P<0.05$], with the same average individual accuracy as seen in the prior studies. When we examined the majority group competence judgments, comparing candidates who were characterized as competent by a majority of our participants with those who had won elections, the association trended positively [55% , $\chi^2(1)=1.00$, $P>0.1$, against an expected 50%] but did not reach statistical significance. This is likely due to our fMRI-scale sample size, as sample sizes of 40 or more are generally required to achieve reliability on this particular measure for competence judgments (Todorov *et al.*, 2005).

Consistent with the robust effect of election loss we found in our first study, a novel behavioral finding from the second study was that the strongest association between election outcome and trait judgments was seen for personal threat judgments. Majority group personal threat judgments corresponded to election loss 65% of the time [$\chi^2(1)=9.00$, $P<0.05$], and average individual accuracy was also above chance [$57 \pm 10\%$, $t(15)=2.65$, $P<0.05$]. In fact, the association between majority personal threat judgments and election outcome was stronger than that for competence [$\chi^2(1)=4.04$, $P<0.05$] and public deceitfulness [$\chi^2(1)=9.00$, $P<0.05$], although not reliably different from attractiveness [$\chi^2(1)=2.01$, $P>0.1$]. In addition, only the association between personal threat judgments and election loss survived in a multiple binomial regression model relating all four social judgments to the election outcomes (beta = 1.5, $P=0.03$, $r^2=0.1$, $P=0.01$).

As previously reported (Todorov *et al.*, 2005), the association between election outcome and attractiveness judgments was not statistically different from chance for the average individual [$46 \pm 10\%$, $t(15)=-1.46$, $P>0.1$] or for majority group judgments [58% correspondence with election loss, $\chi^2(1)=2.56$, $P>0.1$]. Judgments of public deceitfulness across individuals [$49 \pm 9\%$, $t(15)=-0.34$, $P>0.1$] and group majority judgments [50%, ($\chi^2(1)=0$, $P>0.1$)] also did not differ from chance in associating with election outcome.

Our behavioral findings from the second study are, thus, consistent with what we inferred from our first study: there appears to be a primary role for negative attributions in mediating the effect of candidate appearance on election outcome. Interestingly, this may be especially the case for

attributions that affect one's personal welfare (i.e. the personal threat judgments viewers made). Given these behavioral findings, and given that our first study revealed no activations for winners, we focus subsequent imaging analyses on the condition of personal threat (see Supporting Information for more detail). This allows us to determine if loser-elicited activations seen in our first study are also seen here for candidates who lost real elections, under conditions where we aimed to most enhance negative attribution (i.e. in judging personal threat from a smiling politician).

We analyzed the fMRI data from our second study by estimating a general linear model in which separate regressors were formed for the first onset of images based on whether those candidates had won or lost in real elections and on whether those candidates were the majority choice with respect to reflecting a particular trait (see Methods for details). We contrasted the parameter estimates obtained in response to the pictures of candidates who had won and those who had lost real elections. We report significant activations surviving FWE-corrected thresholding at $P < 0.05$ (see Supporting Information).

Again consistent with Study 1, we found no significant activations in our regions of interest for candidates who won real elections (see Supporting Information for complete details). Instead, we found that candidates who lost real elections, compared to those who won, elicited greater activation in the insula/parainsula [18 voxels (45, 0, -15), $t = 4.80$; Figure 2B and C; Table S2B.] and in the ventral anterior cingulate cortex [24 voxels (15, 39, 0); $t = 4.02$; (9,45,6), $t = 3.90$; Figure 3B; Table S2B.]. These locations are within the regions we found for Study 1 and further support the idea that negative attribution is primary in mediating the effects of candidate appearance on voter decisions, and itself is mediated by a network of structures that include the insula/parainsula and ventral anterior cingulate regions. Further evidence for this interpretation comes from the observation that losers elicited an increase in activation in the insula, while winners elicited a decrease (Figure 2C).

To link our two studies directly, we first chose a region in the right insula from Study 1 (with simulated voting) and queried this region [a 10 mm radius sphere centered on the peak voxel (48, -3, -9)] with respect to the contrast effects seen in Study 2 (with real voting). Consistent with Study 1, we found in Study 2 that the contrast of loser > winner, under the condition of threat judgment, resulted in a significantly enhanced activation in this region [$t(15) = 1.87$, $P < 0.05$; Figure 2D]. However, we found no significant effect under any of the other judgment conditions ($P > 0.1$). Similarly, we chose a region in the right anterior cingulate from Study 1 and queried this region [a 10 mm radius sphere centered on the peak voxel (3, 33, 9)] with respect to the contrast effects seen in Study 2. Again, we found that the contrast of loser > winner resulted in a significantly enhanced activation in the region, for threat

judgment only [$t(15) = 1.98$, $P < 0.05$; Figure 3C; all other conditions, $P > 0.1$, except for attractiveness, which shows a near-significant effect of winner > loser, $t(15) = 1.71$, $P = 0.054$]. Thus, brain regions identified in Study 1 showing a differential sensitivity for images of election losers, compared to winners, show this same sensitivity in Study 2 under conditions in which negative attributions are putatively enhanced, and this time for real election outcomes. This is further evidence that negative attributions are primary in mediating the effect of appearance on voting.

While analyses in Study 2 focused primarily on the threat judgment condition, it is important to note that we observed activations under the other three conditions also (see Supporting Information for details). The direction of the significant election contrasts (loser > winner, or winner > loser) was in line with what one would expect given the valence of the social judgment condition. Thus, both threat and deceit conditions produced activations primarily for loser > winner, whereas attractiveness and competence conditions produced activations primarily for winner > loser. We interpret these data to show that either positive or negative attributions can be enhanced with a sufficiently valenced social judgment context, while negative attributions are primary under the context of voting, particularly when there is a lack of other information about the candidates.

DISCUSSION

The activation patterns in the insula and anterior cingulate are similar between our two studies, especially considering that they (i) involved different groups of participants, (ii) used different images of political candidates, (iii) used different tasks, and (iv) used different measures of electoral outcome (i.e. simulated and real). In both studies, the activations were elicited by images of candidates who had lost in an election, simulated or real, consistent with the notion that the activations reflect processing in negative attribution. Taken together, the studies suggest that elicitation of negative emotional processes may predominate in mediating the connection between candidate appearance and voting behavior. This interpretation of the data is based on several observations. First, winners of our simulated election elicited no activations in any brain region, while losers elicited robust activation in both the insula and ventral anterior cingulate. While both of these regions have been shown to be sensitive to positive as well as negative aspects of appearance, under various conditions, our interpretation is that here they were responding to negative aspects since they were strongly activated by candidates who lost. This is consistent with literature associating the insula/parainsula (Coan *et al.*, 2006; Lamm *et al.*, 2007) and the ventral anterior cingulate (Eisenberger *et al.*, 2003; Somerville *et al.*, 2006) with the processing of negatively valenced emotions in social situations. The insula is an area known to mediate interoceptive processing and feelings (Craig, 2002), such as sensations of pain or internal discomfort (Singer *et al.*, 2004; Coan *et al.*,

2006), and the right ventral anterior cingulate is implicated in panic attacks (Eser *et al.*, in press), fear (Williams *et al.*, 2006; Bryant *et al.*, 2007) and uncontrollable pain (Salomons *et al.*, 2007). These regions, including the parainsula (Stefanacci and Amaral, 2002), are also known to connect strongly with the amygdala (Amaral *et al.*, 1992), a structure known to play a key role in negative affect associated with faces. Second, judgments of personal threat were most robustly correlated with election outcomes. Finally, under conditions likely to enhance negative attributions (i.e. examining faces for personal threat), we again saw that candidates who won elections elicited no activations in our regions of interest, while those who lost elections elicited greater activation in the insula and ventral anterior cingulate.

These findings are all the more surprising given that nearly all of our politicians were smiling (92% in Study 1; 100% in Study 2) and none showed any overt negative facial expressions. To our knowledge, this is the first demonstration that images with expressions that are overtly positive can nonetheless drive brain activations related to negative evaluations. It is also the first demonstration of a link between voter decisions and brain activations in people making social judgments.

It is important to qualify the findings in this paper in several ways. First, although the weight of our findings suggests a preferential role for negative attributions from candidate appearance, we did see some areas associated with positive emotional processing in Study 2. However, the behavioral judgments under these conditions (competence and attractiveness; cf. Supporting Information) were not as robustly correlated with real electoral outcomes. Thus, we do not wish to rule out that positive attributions may contribute to the effect of appearance on voter decision making, but this effect may be small compared to the effect of negative attributions. It is also possible that positive attributions are simply more variable across individuals than negative attributions, thus diluting their group effect. Nevertheless, our findings support a model in which the contribution made by negative attributions predominates when voters make decisions based on limited information, in line with findings from political science (Lau and Pomper, 2001; Martin, 2004; Stevens *et al.*, 2008).

A fundamental question in politics is the extent to which voters' decisions are driven by positive motives, which induce them to vote for candidates that they like, or by negative ones, which induce them to vote for the candidate that they do *not* dislike (i.e. negative voting). As detailed above, there is evidence that negative motives play a role, if not an exclusive one, in voters' decisions (Kernell, 1977; Lau, 1985; Fiorina and Shepsle, 1989). The results from our two studies suggest that political 'intangibles', such as a candidate's appearance, might also work primarily via negative motives. This raises a final question about the nature of those intangibles: what is it about a person's appearance that signals negative traits and influences election loss?

Future studies with considerably larger stimulus sets, and with experimental manipulations of facial features, will be required to address this question.

METHODS

Study 1: simulated vote

Participants. Twenty-four participants (seven female, aged 18–38) participated in the study. Participants had no history of neurological or psychiatric illness and were not on psychotropic medications. Participants had no previous knowledge of any of the political candidates whose images were used in the study, and reported no recognition of any of the politicians. All procedures were approved by the Institutional Review Board at the California Institute of Technology, and participant consent was obtained according to the Declaration of Helsinki.

Stimuli. Stimuli consisted of 200 grayscale images of political candidates who ran in the real 2006 U.S. midterm elections for either the Senate (60 images), the House of Representatives (74 images), or Governor (66 images). The stimuli were collected from the candidates' campaign Web sites and other Internet sources. An electoral pair consisted of two images of candidates, one Republican and one Democrat, who ran against one another in the real election. Due to the racial and gender composition of the candidates, 70 of the 100 pairs were of male politicians, and 88 of 100 pairs involved two Caucasian politicians. An independent observer classified 92% of the images as 'smiling'. In 57% of the pairs, both candidates were frontal facing; in the rest at least one was facing to the side. Except for transforming color images into a gray scale, the stimuli were not modified. Images were presented using video goggles (Resonance Technologies Inc.; <http://www.mrivideo.com>). The stimulus presentation and response recording was controlled by Cogent 2000 (Wellcome Department of Imaging Neuroscience; <http://www.vislab.ucl.ac.uk/Cogent/index.html>).

The study was conducted in the month before the 2006 election. An effort was made to avoid pairs in which one of the candidates (e.g. Hillary Clinton) had national prominence or participated in a California election, and familiarity ratings collected from all of the participants after the scanning task verified the stimuli were unfamiliar. On a scale of 1 (completely unfamiliar) to 7 (very familiar), the mean rating was 1.65 (s.d. = 1.03) for the Democratic candidates and 1.62 (s.d. = 0.98) for the Republican candidates.

Procedure. Participants were instructed that they would be asked to vote for real political candidates who were running against each other in the upcoming midterm election. In particular, they were asked to decide who they would be more likely to vote for given that the only information that they had about the politicians were their portraits.

Each trial consisted of three events (Figure 1A). First, a picture of one of the candidates was centrally presented for 1 s. Second, after a blank screen of length 1–10 s (uniform distribution), the picture of the other candidate in the pair

was presented for 1 s. Third, after another blank screen of length 1–10 s, the pictures of both candidates were presented side by side. At this point, participants were asked to cast their vote by pressing either the left or the right button. They had a maximum of 2 s to make a decision. Participants made a response within this time frame in 100% of the trials. Trials were separated by a 1–10 s blank screen. The order of presentation of the candidates as well as their position in the final screen was fully randomized between participants.

Neuroimaging data acquisition. Imaging data were collected on a Siemens 3.0-T Trio MRI scanner. Whole-brain, high-resolution ($1 \times 1 \times 1 \text{ mm}^3$) T1-weighted images were collected for each participant and coregistered with the mean functional, T2*-weighted images. For the fMRI data, we collected gradient-echo T2*-weighted echoplanar images with blood oxygenation level-dependent contrast using an interleaved, ascending image sequence (parameters: TR = 2.75 s, TE = 30 ms, field of view = 192 mm, 44 slices at 3 mm thick, 64×64 voxels, resulting in a voxel size of $3 \times 3 \times 3 \text{ mm}^3$). We used a tilted acquisition, at 30° relative to the anterior commissure–posterior commissure line in order to achieve good signal in both the orbitofrontal cortex and subcortical regions. In addition, we used an eight-channel phased array coil that yields a 40% signal increase in OFC over the standard head coil.

Neuroimaging data analysis. All data analysis was performed using SPM5. We discarded the first five EPI images to allow for signal equilibration, applied slice-timing correction (centered at TR/2), realigned all volumes, spatially normalized a standard EPI template with a resampled isotropic voxel size of $3 \times 3 \times 3 \text{ mm}^3$, spatially smoothed the data using a Gaussian kernel with FWHM of 8 mm and applied intensity normalization and high-pass temporal filtering (filter width 128 s).

We constructed regressors corresponding to the onsets of the images and additionally used six motion regressors. We used the lab vote share as a parametric modulator of the image onset regressor. The lab vote share is the fraction of our participant group who voted for the politician, and the negative lab vote share is simply this fraction subtracted from 1. In our first analysis, we used the positive lab vote share as a parametric modulator, and in our second analysis, we used the negative lab vote share as a parametric modulator. We applied linear contrasts within each participant, and took this to the random effects level using *t*-tests. To select statistically significant clusters, we applied an FWE-corrected threshold of $P < 0.05$ (see Supporting Information for details).

Study 2: trait judgments

Participants. Twenty-two Caucasian women (aged 20–35) participated in the study (note: none of these individuals participated in Study 1). At the time of the experiment, they were registered to vote and had voted in one or more of the following national elections: 2000, 2002 and/or 2004.

Participants had no history of neurological or psychiatric illness and were not on psychoactive medications. Participants had no prior knowledge of any of the political candidates whose images were used in the study and reported no recognition of any of the politicians. Neuroimaging data from six participants were rejected due to excessive motion. The behavioral data of the 16 participants included in the study did not differ significantly from those of six participants rejected for excessive motion. All procedures were approved by the Institutional Review Board at the California Institute of Technology.

Stimuli. Stimuli consisted of 60 grayscale images of smiling political candidates who ran in real U.S. elections for the House of Representatives or Senate in either 2000, 2002 or 2004 (30 pairs of opponents). The stimuli comprised a subset of those used in the 2005 study by Todorov *et al.* (2005), and were selected by three of the experimenters so that both images in an electoral pair (i) were frontal facing, (ii) were of the same gender and ethnicity and (iii) had clear, approximately central presentation of faces that were of approximately the same size. An electoral pair consisted of two images of candidates, one Republican and one Democrat, who actually ran against one another in a real election. Due to the racial/ethnic and gender composition of the original image library, all stimuli were of Caucasian politicians, and 8 of the 30 pairs were of female politicians. Stimuli were preprocessed to normalize overall image intensity while maintaining good image quality, across all 60 images. All images were presented centrally, via an LCD projector and a rear-projection screen, onto a mirror attached to the MRI head coil, approximately 10 inches from a participant's eyes. Stimuli subtended approximately 8° of visual angle. Stimulus control and response recording used the Psychophysics Toolbox v2.54 (Brainard, 1997; Pelli, 1997) in Matlab (the Mathworks, Natick, MA). A pilot behavioral study confirmed that the social judgments made about our selected stimuli were representative of the entire set of face stimuli from which they were drawn (the entire set used by Todorov *et al.* (2005)).

Procedure. Participants were instructed that they would be asked to make judgments about real political candidates who ran against one another in real elections. They were told that they would only be given the images of the politicians to inform their judgments. Image order was counterbalanced across participants. Participants made judgments about candidates' attractiveness (Attr), competence (Comp), public deceitfulness (Dect) and personal threat (Thrt) in four separate scanning sessions. For threat judgments, participants were asked which candidate in a pair looked more likely to act in a physically threatening manner toward them (i.e. personal threat). For attractiveness judgments, participants were asked which candidate looked more physically attractive to them (i.e. personal attractiveness). For competence judgments, participants were asked which candidate looked more competent to hold national congressional office.

For public deceitfulness, participants were asked which candidate looked more likely to lie to the voters (i.e. public deceit). Each session took approximately 9 min to complete. Of 16 participants, six made personal threat decisions prior to making other decisions, while the remainder made personal threat decisions after making decisions about other attributes. There were no effects of block order.

Each trial in a decision block consisted of the sequential presentation of two images in an electoral pair, image A then image B, until a participant entered a decision about the pair via a button press (Figure 1B). This follows a protocol we have used successfully in prior studies of face preference (Kim *et al.*, 2007). An A/B cycle on a given trial proceeded as follows: (i) central presentation of a fixation rectangle that surrounded the area in which an image was to appear; (ii) after 4–6 s, a 30 ms display of image A surrounded by the fixation box, accompanied by a small black dot in the lower left corner (indicating that this was image A); and (iii) after 3–4 s, a 30 ms display of image B surrounded by the fixation box, accompanied by a small black dot in the lower right corner (indicating that this was image B). Cycles were separated by 4–6 s and continued until a participant entered a button press or until 30 s had elapsed, whichever came first (no participant ever took the 30 s). Participants were asked to attend overtly to the space inside the rectangle in preparation for a candidate image. We used eyetracking (MRI-compatible Long-Range Optics Model, Applied Science Laboratories, Bedford, MA) to ensure that participants were looking at the stimuli. Trials that required just one A/B cycle for participants to make their judgment are referred to as one-cycle trials, those that required two cycles are two-cycle trials, and so on.

Behavioral data analysis. Participants primarily took two cycles to decide [$76 \pm 2\%$ of trials were two-cycle trials (mean \pm s.d. across 16 participants, all judgment conditions)]. There was not enough data from other types of trials to conduct a full random effects analysis, so only data from two-cycle trials were examined.

Correlations between social judgments. For each candidate and for each judgment, we first calculated the judgment share, which was just the proportion of participants who decided that candidate was more threatening, attractive, competent or deceitful. Using these values, we calculated Pearson correlation coefficients between the different social judgments that participants made.

Correspondence between social judgments and electoral outcome. We conducted two types of analyses, as done for previous studies (Todorov *et al.*, 2005). To determine average individual association between judgments and real electoral outcome, we calculated the percent agreement between each participant's social judgments and electoral outcome, and then averaged across participants. A simple *t*-test allowed us to determine whether the mean individual association differed significantly from chance (50%). To determine the association between majority group judgment and real

electoral outcome, we first counted which politician in a pair had the most participants naming them as attractive, competent, deceitful and threatening. We then calculated the percent agreement between this social judgment outcome and the signed electoral outcome (for instance, agreement between who was judged to look more competent and real-world election winning; or who was judged to look more threatening and real-world election losing). A χ^2 test finally determined whether majority associations were different from chance (50%).

Neuroimaging data acquisition. This was as for Study 1, except that TR = 2 s, and each EPI image had 34 slices at 4 mm thick.

Neuroimaging data analysis. All data preprocessing was done in SPM2 and analyses were conducted with SPM5 (Wellcome Department of Imaging Neuroscience, London, UK; <http://www.fil.ion.ucl.ac.uk/spm/>) (Friston *et al.*, 2005). Preprocessing was identical to Study 1, except that we used a spatial smoothing filter of 6 mm FWHM.

We combined all four judgment conditions in one design matrix and proceeded in three steps. In the first step, we estimated a general linear model with AR(1). Second, we calculated first-level contrasts of group chosen (GC) vs group unchosen (GU) for each of the four judgment choices, and real-world election winner (RW) vs loser (RL), again for each of the four judgment conditions, as well as interactions between conditions. Finally, for each of these first-level contrasts, we calculated a second-level random effects contrast using a one sample *t*-test. To select statistically significant clusters, we applied an FWE-corrected threshold of $P < 0.05$ (see Supporting Information for details). Our analysis focused on the first cycle of the 2-cycle trials. We did this for two reasons: to maintain consistency with our first study and to maximize thin-slice conditions.

To examine second-level effects in Study 2 using activation ROIs from Study 1, we used the RFXPLOT toolbox for SPM5 (<http://rfxplot.sourceforge.net/>). We formed a mask to select all voxels inside a sphere (10 mm radius) for a given ROI from Study 1 [right insula: around (48, -3, -9); right anterior cingulate: around (3, 33, 9)]. We applied this mask to the individual first-level contrast images from participants in Study 2, selecting all voxels inside the mask (both sub- and suprathreshold) and calculated a mean ROI contrast for each participant. We used *t*-tests to determine whether the group mean contrasts were significantly different from zero.

SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

REFERENCES

- Alvarez, R.M. (1997). *Information and Elections*. Ann Arbor, MI: University of Michigan, pp. 109–56.
- Amaral, D.G., Price, J.L., Pitkänen, A., Carmichael, S.T. (1992). The anatomical organization of the primate amygdaloid complex. In: Aggleton, J.P.,

- editor. *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*. New York: Wiley-Liss, pp. 1–67.
- Bailenson, J.N., Garland, P., Iyengar, S., Yee, N. (2006). Transformed facial similarity as a political cue: a preliminary investigation. *Political Psychology*, 27, 373–85.
- Ballew, C.C., Todorov, A. (2007). Predicting political elections from rapid and unreflective face decisions. *Proceedings of the National Academy of Science*, 104, 17948–53.
- Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–6.
- Bryant, R.A., Felmingham, K., Kemp, A., et al. (2007). Amygdala and ventral anterior cingulate activation predicts treatment response to cognitive behaviour therapy for post-traumatic stress disorder. *Psychological Medicine*, 38, 556–61.
- CNN (2006). Record amount spent on campaign ads in 2006. <http://www.cnn.com/POLITICS/blogs/politicalticker/2006/12/record-amount-spent-on-campaign-ads-in.html>.
- Coan, J.A., Schaefer, H.S., Davidson, R.J. (2006). Lending a hand: social regulation of the neural response to threat. *Psychological Science*, 17, 1032–9.
- Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29, 162–173.
- Craig, A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Natural Reviews Neuroscience*, 3, 655–66.
- Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper, pp. 207–59.
- Eisenberger, N.I., Lieberman, M.D., Williams, K.D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, 302, 290–2.
- Eser, D., Leicht, G., Lutz, J., et al. (in press). Functional neuroanatomy of CCK-4-induced panic attacks in healthy volunteers. *Human Brain Mapping*. [Epub ahead of print; 19 December 2007].
- Fiorina, M.P., Shepsle, K.A. (1989). Is negative voting an artifact? *American Journal of Political Science*, 33, 423–39.
- Forgas, J.P., Fiedler, K. (1996). Us and them: mood effects on intergroup discrimination. *Journal of Personality and Social Psychology*, 70, 28–40.
- Friston, K.J., Williams, S.C., Howard, R., Frackowiak, R.S.J., Turner, R. (2005). Online resource, Wellcome Trust Centre for Neuroimaging. <http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>.
- Galdi, S., Arcuri, L., Gawronski, B. (2008). Automatic mental associations predict future choices of undecided decision-makers. *Science*, 321, 1100–02.
- Gottfried, J.A., O'Doherty, J., Dolan, R.J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science*, 301, 1104–07.
- Hassin, R., Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, 78, 837–52.
- Iyengar, S., Simon, A.F. (2000). New perspectives and evidence on political communication and campaign effects. *Annual Review of Psychology*, 51, 149–69.
- Kanwisher, N., Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 361, 2109–28.
- Kernell, S. (1977). Presidential popularity and negative voting. *American Political Science Review*, 71, 44–66.
- Kim, H., Adolphs, R., O'Doherty, J.P., Shimojo, S. (2007). Temporal isolation of neural processes underlying face preference decisions. *Proceedings of the National Academy of Sciences of the USA*, 104, 18253–58.
- Lamm, C., Batson, C.D., Decety, J. (2007). The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, 19, 42–58.
- Lau, R.R. (1982). Negativity in political perception. *Political Behavior*, 4, 353–78.
- Lau, R.R. (1985). Two explanations for negativity effects in political behavior. *American Journal of Political Science*, 29, 119–138.
- Lau, R.R., Pomper, G.M. (2001). Effects of negative campaigning on turnout in U.S. Senate elections, 1988–1998. *Journal of Politics*, 63, 804–19.
- Martin, P.S. (2004). Inside the black box of negative campaign effects: three reasons why negative campaigns mobilize. *Political Psychology*, 25, 545–562.
- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K.J., Dolan, R.J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304, 452–4.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10, 437–42.
- Ploughman, A., Tracey, I., Gati, J.S., et al. (1999). Dissociating pain from its anticipation in the human brain. *Science*, 284, 1979–81.
- Popkin, S. (1991) *The Reasoning Voter: Communication and Persuasions in Presidential Campaigns*. Chicago: University of Chicago Press, pp. 7–43.
- Prior, M. (2005). News vs. entertainment: how increasing media choice widens gaps in political knowledge and turnout. *American Journal of Political Science*, 49, 577.
- Rosenberg, S.W. (1986). The image and the vote: the effect of candidate presentation on voter preferences. *American Journal of Political Science*, 30, 108–27.
- Rosenberg, S.W., Kahn, S., Tran, T., Le, M.-T. (1991). Creating a political image: shaping appearance and manipulating the vote. *Political Behavior*, 13, 345–67.
- Salomons, T.V., Johnstone, T., Backonja, M.M., Shackman, A.J., Davidson, R.J. (2007). Individual differences in the effects of perceived controllability on pain perception: critical role of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 19, 993–1003.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303, 1157–62.
- Somerville, L.H., Heatherton, T.F., Kelley, W.M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, 9, 1007–8.
- Stefanacci, L., Amaral, D.G. (2002). Some observations on cortical inputs to the macaque monkey amygdala: an anterograde tracing study. *Journal of Comparative Neurology*, 451, 301–23.
- Stevens, D., Sullivan, J., Allen, B., Alger, D. (2008). What's good for the goose is bad for the gander: negative political advertising, partisanship, and turnout. *Journal of Politics*, 70, 527–41.
- Todorov, A., Mandisodza, A.N., Goren, A., Hall, C.C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308, 1623–26.
- Todorov, A., Uleman, J.S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, 39, 549–62.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15, 273–89.
- Williams, L.M., Liddell, B.J., Kemp, A.H., et al. (2006). Amygdala-prefrontal dissociation of subliminal and supraliminal fear. *Human Brain Mapping*, 27, 652–61.
- Willis, J., Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, 592–8.
- Winston, J.S., Strange, B.A., O'Doherty, J., Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, 5, 277–83.
- Xiong, J., Gao, J.-H., Lancaster, J.L., Fox, P.T. (1995). Clustered analysis for functional MRI activation studies of the human brain. *Human Brain Mapping*, 3, 287–301.

LETTERS

Oxytocin increases trust in humans

Michael Kosfeld^{1*}, Markus Heinrichs^{2*}, Paul J. Zak³, Urs Fischbacher¹ & Ernst Fehr^{1,4}

Trust pervades human societies^{1,2}. Trust is indispensable in friendship, love, families and organizations, and plays a key role in economic exchange and politics³. In the absence of trust among trading partners, market transactions break down. In the absence of trust in a country's institutions and leaders, political legitimacy breaks down. Much recent evidence indicates that trust contributes to economic, political and social success^{4,5}. Little is known, however, about the biological basis of trust among humans. Here we show that intranasal administration of oxytocin, a neuropeptide that plays a key role in social attachment and affiliation in non-human mammals^{6–8}, causes a substantial increase in trust among humans, thereby greatly increasing the benefits from social interactions. We also show that the effect of oxytocin on trust is not due to a general increase in the readiness to bear risks. On the contrary, oxytocin specifically affects an individual's willingness to accept social risks arising through interpersonal interactions. These results concur with animal research suggesting an essential role for oxytocin as a biological basis of prosocial approach behaviour.

In non-human mammals, the neuropeptide oxytocin has a central role in general behavioural regulation, particularly in positive social interactions. Aside from its well-known physiological functions in milk letdown and during labour, oxytocin receptors are distributed in various brain regions associated with behaviour^{9,10}, including pair bonding, maternal care, sexual behaviour, and the ability to form normal social attachments^{6–8,11–15}. Thus, oxytocin seems to permit animals to overcome their natural avoidance of proximity and thereby facilitates approach behaviour. Given that oxytocin is believed to promote social attachment and affiliation in non-human mammals, we hypothesized that oxytocin might also promote prosocial approach behaviours—such as trust—in humans. Recent research has shown that neuropeptides cross the blood-brain barrier after intranasal administration¹⁶, providing a useful method for studying the central nervous system effects of oxytocin in humans^{17,18}. We used a double-blind study design to compare trusting behaviour in a group of subjects that received a single dose of intranasal oxytocin with that of subjects in a control group that received placebo.

We analysed the effect of exogenously administered oxytocin on individuals' decisions in a trust game with real monetary stakes^{19–22}. In this trust game, two subjects interacting anonymously play either the role of an investor or a trustee (Fig. 1). First, the investor has the option of choosing a costly trusting action by giving money to the trustee. If the investor transfers money, the total amount available for distribution between the two players increases but, initially, the trustee reaps the whole increase. The trustee is then informed about the investor's transfer and can honour the investor's trust by sharing the monetary increase generated by the investor's transfer. Thus, if the investor gives money to the trustee and the latter shares the proceeds of the transfer, both players end up with a higher

monetary payoff. However, the trustee also has the option of violating the investor's trust. As sharing the proceeds is costly for the trustee, a selfish trustee will never honour the investor's trust because the investor and the trustee interact only once during the experiment.

The investor is therefore caught in a dilemma: if he trusts and the trustee shares, the investor increases his payoff, but he is also subject to the risk that the trustee will abuse this trust. In the latter case, the investor is worse off than if he had not trusted at all and, adding insult to injury, the trustee has an unfair payoff advantage relative to the investor. Substantial evidence exists to show that humans are averse to such risks^{22–24}. Moreover, the aversion of investors to abuse of trust seems to have an important role across different human cultures and social groups in the context of our game^{22,25}. The investors have to overcome their aversion against these risks in order to trust, allowing us to address the question of whether oxytocin modulates this trusting behaviour in humans.

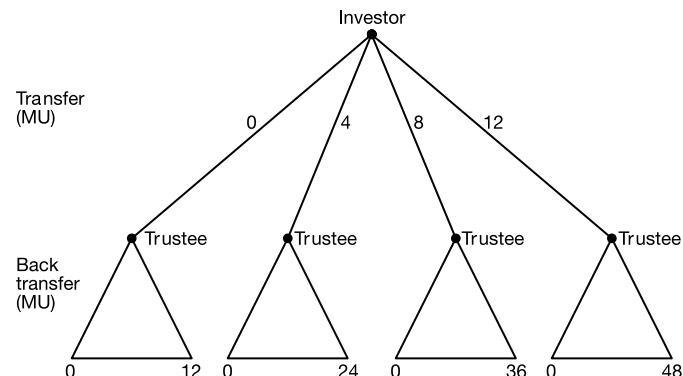


Figure 1 | The trust game. Both subjects receive an initial endowment of 12 monetary units (MU). The investor can send 0, 4, 8 or 12 MU to the trustee. The experimenter triples each MU the investor transfers. After the investor's decision is made, the trustee is informed about the investor's transfer. Then the trustee has the option of sending any amount between zero and his total amount available back to the investor. For example, if the investor has sent 12 MU, the trustee possesses 48 MU (12 MU own endowment + 36 MU tripled transfer) and can, therefore choose any back transfer from 0 to 48 MUs. The experimenter does not triple the back transfer. The investor's final payoff corresponds to the initial endowment minus the transfer to the trustee, plus the back transfer from the trustee. The trustee's final payoff is given by his initial endowment plus the tripled transfer of the investor, minus the back transfer to the investor. At the end of the experiment, the earned MU are exchanged into real money according to a publicly announced exchange rate (see Methods). Each subject made four decisions in the same player role while paired with four different, randomly selected interaction partners.

¹University of Zurich, Institute for Empirical Research in Economics, Blumlisalpstrasse 10, CH-8006 Zurich, Switzerland. ²University of Zurich, Department of Clinical Psychology and Psychotherapy, Zurichbergstrasse 43, CH-8044 Zurich, Switzerland. ³Center for Neuroeconomics Studies, Claremont Graduate University, Claremont, California 91711-6165, USA. ⁴Collegium Helveticum, Schmelzbergstrasse 25, CH-8092 Zurich, Switzerland.

*These authors contributed equally to this work.

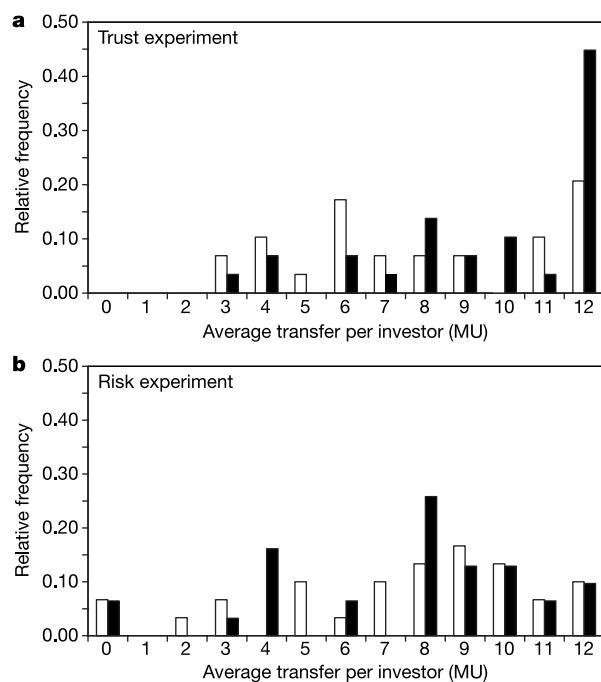


Figure 2 | Transfers in the trust and the risk experiment. Each observation represents the average transfer amount (in MU) over four transfer decisions per investor. **a**, Relative frequency of investors' average transfers in oxytocin (filled bars) and placebo (open bars) groups in the trust experiment ($n = 58$). Subjects given oxytocin show significantly higher transfer levels. **b**, Relative frequency of investors' average transfers in oxytocin (filled bars) and placebo (open bars) groups in the risk experiment ($n = 61$). Subjects in the oxytocin and the placebo group show statistically identical transfer levels.

Our hypothesis that oxytocin increases the trusting behaviour of investors implies that the investors in the oxytocin group ($n = 29$) will show higher money transfers than those in the placebo group ($n = 29$). In fact, our data show that oxytocin increases investors' trust considerably. Out of the 29 subjects, 13 (45%) in the oxytocin group showed the maximal trust level, whereas only 6 of the 29 subjects (21%) in the placebo group showed maximal trust (Fig. 2a). In contrast, only 21% of the subjects in the oxytocin group had a trust level below 8 monetary units (MU), but 45% of the subjects in the control group showed such low levels of trust. These differences in the distribution of trust result in higher average and median trust levels for subjects given oxytocin (Table 1). The investors' average transfer is 17% higher in the oxytocin group (Mann-Whitney U -test; $z = -1.897$, $P = 0.029$, one-sided), and the median transfer in the oxytocin group is 10 MU, compared to a median of only 8 MU for subjects in the placebo group.

In the trust game, the risk on the part of the investor's is due to the uncertainty of the trustee's behaviour—that is, a social interaction with a specific trustee constitutes the risk. This raises the question of whether oxytocin helps humans to overcome a general aversion

against risks or whether oxytocin specifically affects trusting behaviour in social interactions. In order to answer this question, we conducted a risk experiment in which the investor faced the same choices as in the trust game but in which a random mechanism, not the trustee's decision, determined the investor's risk. The random mechanism in the risk experiment replicated the trustees' decisions in the trust experiment. Therefore, the investors faced exactly the same risk as in the trust experiment (see Methods); however, their transfer decisions were not embedded in a social interaction because there were no trustees in the risk experiment.

In this risk experiment, the investors' behaviour does not differ between the oxytocin and the placebo groups (Table 1 and Fig. 2b). The median transfer is 8 MU and the average transfer is 7.5 MU in both groups (Mann-Whitney U -test; $z = 0.022$, $P = 0.983$; two-sided test, $n = 31$ in oxytocin group, $n = 30$ in placebo group). Moreover, there is no significant difference in a comparison of the placebo group in the trust experiment with the oxytocin group and the placebo group in the risk experiment (Kruskal-Wallis test; $\chi^2 = 0.533$, d.f. = 2, $P = 0.766$), with identical median transfers across groups (Table 1). However, if we add the oxytocin group in the trust experiment to these three samples, significant differences are observed (Kruskal-Wallis test; $\chi^2 = 8.610$, d.f. = 3, $P = 0.035$), indicating that only the investors in the oxytocin group of the trust experiment behave differently. Thus, oxytocin increases the investors' transfer levels in the trust experiment but not in the risk experiment. This finding is illustrated by a comparison of Figs 2a and b, which show that only 10% of the subjects with oxytocin choose the maximal transfer level in the risk experiment, whereas 45% choose the maximal level in the trust experiment. Therefore, the differences between the oxytocin group in the trust experiment and the oxytocin group in the risk experiment are highly significant (Mann-Whitney U -test; $z = -2.563$, $P = 0.010$, two-sided), suggesting that oxytocin specifically affects trust in interpersonal interactions.

The risk experiment constitutes a powerful control for the effects of oxytocin on trusting behaviour because everything is kept constant relative to the trust experiment, except that the investors' risk in the risk experiment is not generated through a social interaction. Specifically, all the indirect effects of oxytocin on the state of a subject, such as possible effects on mood or calmness, would be present in both the trust and the risk experiment. Therefore, these potential indirect effects of oxytocin cannot be responsible for the effect of oxytocin on trusting behaviour. Moreover, in order to provide an additional control for non-specific effects that might be associated with oxytocin administration, we explicitly measured mood and calmness before substance administration and 50 min after administration (but before subjects played the trust or the risk game). We used a questionnaire suitable for repeated measures within short periods of time, one that is widely used in neuropharmacological studies in humans²⁶ and correlates with physiological measures¹⁷. There were no statistical differences in the levels of mood and calmness before and after the administration of oxytocin in either the trust or the risk experiment. (Trust experiment: $z = -1.541$, $P = 0.123$ for calmness; $z = 1.452$, $P = 0.146$ for mood; $n = 29$. Risk experiment: $z = 0.620$, $P = 0.535$ for calmness; $z = -0.841$, $P = 0.400$ for mood; $n = 31$; two-sided Wilcoxon signed rank tests.) This provides further support for our conclusion

Table 1 | Median and average transfer behaviour of investors

	Trust experiment		Risk experiment	
	Oxytocin group	Placebo group	Oxytocin group	Placebo group
Mean average transfer (MU)	9.6	8.1	7.5	7.5
Median average transfer (MU)	10	8	8	8
Standard deviation of transfers (MU)	2.8	3.1	3.3	3.4
Number of observations	29	29	31	30

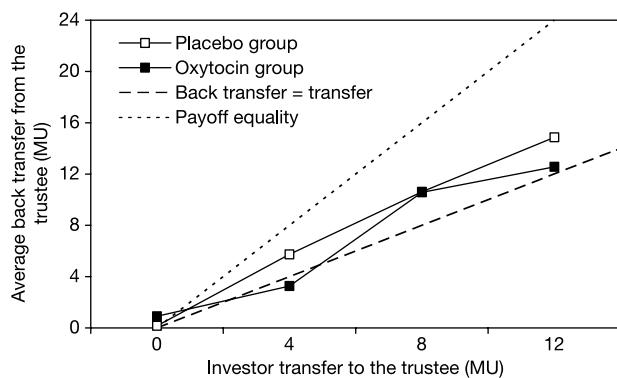


Figure 3 | Average back transfer of trustees to their investors. The graph shows the average back transfer by trustees for different levels of investor transfer in the oxytocin and placebo groups. The dotted line shows the level of the back transfer necessary to achieve payoff equality between the investor and the trustee. The dashed line shows a level of back transfer equal to the investor's transfer to the trustee. The trustees' back transfers are on average slightly higher than the amount sent by the investor. Trustees in both treatment groups make higher back transfers in response to higher original investor transfer levels. However, there is no statistically significant difference in back transfers between subjects in the oxytocin and the placebo groups.

that the effect of oxytocin on human trust is not caused by non-specific, psychotropic effects of oxytocin.

What mechanisms might be involved in generating the effect of oxytocin on trusting behaviour? One possibility is that oxytocin causes a general increase in prosocial inclinations. This implies that oxytocin should affect not only the prosocial behaviour of the investors but also that of the trustees. We would therefore predict that those trustees who are given oxytocin should make higher back transfers at any given level than the trustees who received placebo. However, trustees given oxytocin do not show more trustworthy behaviour (Fig. 3). At every positive transfer level (4, 8 or 12 MU), their back transfers are statistically indistinguishable from those of placebo trustees (Mann Whitney *U*-tests; $P > 0.243$, two-sided tests for each positive transfer level). Thus, oxytocin does not increase the general inclination to behave prosocially. Rather, oxytocin specifically affects the trusting behaviour of investors.

We hypothesize that the differing effect of oxytocin on the behaviour of investors and trustees is related to the fact that investors and trustees face rather different situations. Specifically, investors have to make the first step; they have to 'approach' the trustee by transferring money. In contrast, the trustees can condition their behaviour on the basis of the investors' actions. Thus, the psychology of trust is important for investors, whereas the psychology of strong reciprocity²⁷ is relevant for trustees. The fact that oxytocin affects subjects' approach or trust behaviour, but not their degree of reciprocity, is in agreement with animal studies. There is substantial evidence that oxytocin promotes prosocial approach behaviour by inhibiting defensive behaviours^{6,13}, but there is no evidence that oxytocin affects reciprocity in animals.

A second mechanism behind the effect of oxytocin on trust could be based on subjects' beliefs. Oxytocin might render subjects more optimistic about the likelihood of a good outcome. In order to address this question, we measured the investor's subjective expectation about the trustee's back transfer after every transfer decision. A Mann-Whitney *U*-test indicates that these expectations do not differ significantly between oxytocin and placebo groups at every feasible positive transfer level ($P > 0.357$, two-sided tests at transfer levels of 4, 8 or 12 MU). Thus, the investors given oxytocin show more trusting behaviour but do not hold significantly different beliefs about the trustworthiness of others. Moreover, oxytocin

does not affect investors' beliefs about the likelihood of a good outcome in the risk experiment ($P > 0.128$, two-sided Mann Whitney *U*-tests for transfer levels of 4, 8 or 12 MU).

Finally, there is the possibility that oxytocin helps subjects to overcome their betrayal aversion in social interactions. This explanation is consistent with the differing effects of oxytocin across the trust and the risk experiments, and is further supported by the fact that investors faced a considerable betrayal risk. An increase in the transfer level from 4 or 8 MU to 12 MU decreased the investor's average payoff slightly, whereas it increased the objective risk of very low back transfers by the trustee. However, betrayal aversion alone cannot explain why investors given oxytocin make higher transfers in the trust experiment compared with the risk experiment, because betrayal is impossible in the risk experiment. The higher transfers in the trust experiment can be reconciled with betrayal aversion if one acknowledges that investors' behaviour in the trust experiment is also likely to be driven by the motive to increase the available amount for distribution between the two players²⁸. As this motive cannot operate in the risk experiment, it can only increase transfers levels in the trust experiment. Our interpretation of oxytocin's effect on trust in terms of betrayal aversion may be seen in the light of animal studies indicating that increased availability of oxytocin in the central nervous system facilitates approach behaviour, by linking the overcoming of social avoidance with the activation of brain circuits implicated in reward (for example, the nucleus accumbens)^{12,15}.

The ubiquity of trusting behaviour is perhaps one of the distinguishing features of the human species. An element of trust characterizes almost all human social interactions. Here we have sought to examine the effect of oxytocin on trust in humans. Research in non-human mammals suggests that oxytocin has a key role in social attachment and affiliation. We find that intranasal administration of oxytocin causes a substantial increase in trusting behaviour. Subjects given oxytocin seem better able to overcome trust obstacles such as betrayal aversion. Of course, this finding could be misused to induce trusting behaviours that selfish actors subsequently exploit. However, our findings may also have positive clinical implications for patients with mental disorders that are associated with social dysfunctions (for example, social phobia or autism). In particular, social phobia ranks as the third most common mental health disorder and is characterized by marked social deficits, including persistent fear and avoidance of social interactions. Thus, our results might lead to fertile research on the role of oxytocin in several mental health disorders with major public health significance.

METHODS

Subjects. A total of 194 healthy male students (mean age \pm s.d., 22.0 ± 3.4 yr) from different universities in Zurich participated in the study. The trust experiment had 128 participants, and 66 subjects participated in the risk experiment. Exclusion criteria for participation were significant medical or psychiatric illness, medication, smoking more than 15 cigarettes per day, and drug or alcohol abuse. Subjects were instructed to abstain from food and drink (other than water) for 2 h before the experiment, and from alcohol, smoking and caffeine for 24 h before the experiment. Participants were informed at the time of recruitment that the experiment would evaluate the effects of a hormone on decision making. In total, 16 individuals out of the original sample of 194 were excluded because of incorrect substance administration (7 in the trust experiment, 5 in the risk experiment) or their stated disbelief that the opponent in the trust game was actually a human being (4 participants). The study protocol was approved by the ethics committee of the University of Zurich. All subjects gave written, informed consent before participation.

Substance administration. Subjects received a single intranasal dose of 24 IU oxytocin (Syntocinon-Spray, Novartis; 3 puffs per nostril, each with 4 IU oxytocin) or placebo 50 min before the start of the trust or the risk experiment. Subjects were randomly assigned to the oxytocin or placebo group (double-blind, placebo-controlled study design). In order to avoid any subjective substance effects (for example, olfactory effects) other than those caused by oxytocin, the placebo contained all inactive ingredients except for the neuropeptide.

Behavioural experiment and questionnaires. After substance administration,

subjects completed questionnaires on a computer to measure demographic items and psychological characteristics. Owing to the crucial role of the social environment in triggering behavioural effects of oxytocin (as shown in animal research)^{13,29}, subjects were asked to wait in the rest area while the next part of the experiment was prepared. During this 5-min waiting period, subjects were seated at different tables. Subjects at the same table could talk to each other, but at the beginning of the experiment they were informed that they would not be interacting with those subjects who sat at the same table. When subjects re-entered the laboratory for both experiments, they received written instructions (available from the authors on request) explaining the payoff structure of the experiment and the private payment procedure at the end of the experiment. Subjects were randomly and anonymously assigned to the role of investor or trustee in the trust experiment, and did not know the identity of the persons with whom they were matched. After subjects had read the instructions in each experiment, we checked whether they understood the payoff structure by means of several hypothetical examples. All subjects (with one exception) answered the control questions correctly. One subject did not answer the control questions correctly and was excluded from the data set (this subject also did not apply the substance correctly). In addition, subjects received an oral summary of the instructions.

Each subject in the trust experiment made four decisions in the same player role while paired with different, randomly selected interaction partners. No pair of subjects interacted twice. Subjects in the role of the investor received no feedback about the trustee's decision between the different interactions. After every transfer decision, each investor was asked about his belief with regard to the expected back transfer from the trustee. Notably, trust levels were statistically constant across the four decisions. There is no time trend in investors' decisions in either the oxytocin or the placebo group. In the risk experiment, everything was identical to the trust experiment, except that all subjects played the role of an investor who could transfer 0, 4, 8, or 12 MU into a project rather than to a trustee. In particular, an investor's payoff risk (that is, the distribution of payoffs) in the risk experiment was identical to that in the trust experiment at any feasible transfer level.

To measure alterations in the psychological state of subjects throughout the course of the experiment, we assessed their mood and calmness at the beginning of the experiment (before substance administration) and immediately before the trust experiment or the risk experiment, by means of a suitable questionnaire²⁶. All decisions in the experiments and the answers to the questionnaires were entered on a computer using z-Tree software³⁰. Subjects received a flat fee of 80 Swiss francs for participation in the experiment; each MU earned in the trust and the risk experiment was worth 0.40 Swiss francs.

Received 20 April; accepted 5 May 2005.

1. Luhmann, N. *Trust and Power* (Wiley, New York, 1979).
2. Coleman, J. *Foundations of Social Theory* 91–118 (Belknap Press of Harvard Univ. Press, Cambridge, 1990).
3. Arrow, K. Gifts and exchanges. *Phil. Publ. Affairs* 1, 343–362 (1972).
4. Knack, S. & Keefer, P. Does social capital have an economic payoff? A cross-country investigation. *Q. J. Econ.* 112, 1251–1288 (1997).
5. Zak, P. J. & Knack, S. Trust and growth. *Econ. J.* 111, 295–321 (2001).
6. Carter, C. S. Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology* 23, 779–818 (1998).
7. Uvnäs-Moberg, K. Oxytocin may mediate the benefits of positive social interaction and emotions. *Psychoneuroendocrinology* 23, 819–835 (1998).
8. Insel, T. R. & Young, L. J. The neurobiology of attachment. *Nature Rev. Neurosci.* 2, 129–136 (2001).
9. Landgraf, R. & Neumann, I. D. Vasopressin and oxytocin release within the brain: a dynamic concept of multiple and variable modes of neuropeptide communication. *Front. Neuroendocrinol.* 25, 150–176 (2004).

10. Huber, D., Pierre, V. & Ron, S. Vasopressin and oxytocin excite distinct neuronal populations in the central amygdala. *Science* 308, 245–248 (2005).
11. Carter, C. S., Altemus, M. & Chrousos, G. P. Neuroendocrine and emotional changes in the post-partum period. *Prog. Brain Res.* 133, 241–249 (2001).
12. Young, L. J., Lim, M. M., Gingrich, B. & Insel, T. R. Cellular mechanisms of social attachment. *Horm. Behav.* 40, 133–138 (2001).
13. Pedersen, C. A. Oxytocin control of maternal behavior. Regulation by sex steroids and offspring stimuli. *Ann. NY Acad. Sci.* 807, 126–145 (1997).
14. Heinrichs, M., Neumann, I. & Ehlert, U. Lactation and stress: protective effects of breast-feeding in humans. *Stress* 5, 195–203 (2002).
15. Insel, T. R. & Shapiro, L. E. Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proc. Natl. Acad. Sci. USA* 89, 5981–5985 (1992).
16. Born, J. et al. Sniffing neuropeptides: a transnasal approach to the human brain. *Nature Neurosci.* 5, 514–516 (2002).
17. Heinrichs, M., Baumgartner, T., Kirschbaum, C. & Ehlert, U. Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biol. Psychiatry* 54, 1389–1398 (2003).
18. Heinrichs, M., Meinschmidt, G., Wippich, W., Ehlert, U. & Hellhammer, D. H. Selective amnesia effects of oxytocin on human memory. *Physiol. Behav.* 83, 31–38 (2004).
19. Camerer, C. & Weigelt, K. Experimental tests of a sequential equilibrium reputation model. *Econometrica* 56, 1–36 (1988).
20. Fehr, E., Kirchsteiger, G. & Riedl, A. Does fairness prevent market clearing? An experimental investigation. *Q. J. Econ.* 108, 437–459 (1993).
21. Berg, J., Dickhaut, J. & McCabe, K. Trust, reciprocity and social history. *Games Econ. Behav.* 10, 122–142 (1995).
22. Bohnet, I. & Zeckhauser, R. Trust, risk and betrayal. *J. Econ. Behav. Organ.* 55, 467–484 (2004).
23. Holt, C. & Laury, S. Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655 (2002).
24. Fehr, E. & Schmidt, K. M. A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114, 817–868 (1999).
25. Hong, K. & Bohnet, I. *Status and Distrust: the Relevance of Inequality and Betrayal Aversion* (Working Paper RW04-041, Kennedy School, Harvard Univ., Cambridge, 2004).
26. Steyer, R., Schwenkmezger, P., Notz, P. & Eid, M. *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF)* [Multidimensional mood questionnaire] (Hogrefe, Göttingen, 1997).
27. Gintis, H., Bowles, S., Boyd, R. & Fehr, E. Explaining altruistic behavior in humans. *Evol. Hum. Behav.* 24, 153–172 (2003).
28. Engelmann, D. & Strobel, M. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *Am. Econ. Rev.* 94, 857–869 (2004).
29. Kendrick, K. M. et al. Neural control of maternal behaviour and olfactory recognition of offspring. *Brain Res. Bull.* 44, 383–395 (1997).
30. Fischbacher, U. *z-Tree: Zurich Toolbox for Ready-made Economic Experiments* (Working Paper No. 21, Institute for Empirical Research in Economics, Univ., Zurich, 1999).

Acknowledgements This paper is part of the Research Priority Program 'Foundations of Human Social Behaviour—Altruism versus Egoism' at the University of Zurich. Financial support from the MacArthur Foundation (Network on Economic Environments and the Evolution of Individual Preferences and Social Norms) and the Cogito Foundation is gratefully acknowledged. M.H. also acknowledges support from the Swiss National Science Foundation. We thank F. Heusi, A. Enzler, S. Gilomen, E. Götz, M. Koenigsberg, K. Korsunsky, S. Krammer, S. Lauber, P. Obrist-Rybar, M. Schellhammer, B. Studer, R. Stünzi and M. Trottmann for research assistance.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.F. (efehr@iew.unizh.ch) and M.H. (m.heinrichs@psychologie.unizh.ch).

CONSUMER ABERRATIONS



One approach to understand consumer choice is to study cases in which those behaviors go awry. What happens to a consumer when he starts gambling, or what happens in compulsive buying disorder (“shopaholics”)? Besides allowing us to better understand and treat those disorders, this knowledge also allows us to better understand the mechanisms of choice in normal consumer behavior.

Here, we focus on a few selected aberrant consumer behaviors that highlight some of the basic mechanisms of choice. This also provides an important and often forgotten aspect of consumer neuroscience and neuromarketing: the knowledge and insights we are gaining can be used to help sufferers in exactly the kinds of daily situations that things go wrong.

Finally, the aspect of aberrant consumer behaviors also suggests that consumer behavior and marketing can be used as an important and viable path to understand the basic brain mechanisms of choice. Thus, cognitive neuroscientists should look to consumer behavior with a more keen interest than we have seen until now.

Is Compulsive Buying a Real Disorder, and Is It Really Compulsive?

The article in this issue by Koran et al. raises several intriguing questions regarding a novel proposed psychiatric disorder: compulsive buying. DSM provides a working model of categories and diagnostic criteria for psychiatric disorders. DSM is constantly evolving and research planning is underway for DSM-V. Changes to DSM-V being considered include the creation of two broad new categories that may influence the conceptualization of compulsive buying.

A category related to obsessive-compulsive-related disorders might include disorders such as obsessive compulsive disorder, obsessive compulsive personality disorder, hoarding, body dysmorphic disorder, eating disorders, hypochondriasis, Tourette's syndrome, Sydenham's chorea or pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections, and pathological grooming disorders, such as trichotillomania, skin picking, and nail biting. Compulsive buying was not determined to be a good fit for this category. On the other hand, a parallel category under consideration is behavioral and substance addictions, which might include substance-related disorders and several impulse-control disorders (pathological gambling, pyromania, and kleptomania), as well as others currently in the category of impulse control disorders not otherwise specified (Internet addiction, impulsive-compulsive sexual behavior, and compulsive buying). The National Institute on Drug Abuse has considered behavioral addictions (such as compulsive buying) to be "cleaner" and more homogeneous models of substance addictions because these conditions may share clinical features and perhaps underlying brain circuitry, and these features and circuitry are not altered by the ingestion of exogenous substances. Similar phases seem to occur for behavioral and substance addictions: initially, episodes are characterized by increasing physiological and emotional arousal before the act; pleasure, high, or gratification associated with the act; and a decrease in arousal and feelings of guilt and remorse afterward. Tolerance and physiological withdrawal can also develop. Because an impulsive component (pleasure, arousal, or gratification) is involved in initiating the cycle, and a compulsive component is involved in the persistence of the behavior, these conditions may also be thought of as impulsive-compulsive disorders.

The creation of a condition such as compulsive buying might be associated with controversy and criticized by some as creating a trivial disorder; "medicalizing" a "moral" problem or creating a new disorder in order to sell more pharmaceuticals. Similar criticisms of attention deficit hyperactivity disorder (ADHD) and social anxiety disorder have been raised: that children with minor and natural levels of excess activity should not be "medicalized" or medicated or that because so many people are socially anxious, this is a natural trait not worthy of diagnosis or treatment. However, the issues involved in creating new diagnoses is complex.

In this issue, Koran et al. reported on a study of compulsive buying. They surveyed a large random sample of U.S. adults to estimate a prevalence rate and to characterize compulsive buyers. They and others have proposed names and diagnostic criteria for this problem and, as required for most DSM disorders, the criteria include significant

"Changes to DSM-V being considered include the creation of two broad new categories that may influence the conceptualization of compulsive buying."

distress or functional impairment, as well as criteria specific to the disorder. As is typical at this stage, the specific name and criteria differ from researcher to researcher and study to study, complicating the development of knowledge about the condition. Until a certain amount of evidence of a new disorder is accumulated, not enough is known to define criteria, but at a certain point, there is enough information to propose criteria. Including a disorder in DSM is very helpful for the advance of knowledge because researchers can then use the defined criteria in their new research, and the criteria can be refined over time as more research is completed.

Clearly, the behavioral addictions or impulse control disorders can be viewed from different perspectives, including: a medical perspective; a moral, ethical, or religious perspective; and a legal perspective. These behaviors exist on a continuum, perhaps in a normal distribution in the general population, with many individuals having some of the behaviors, a few showing none, and a few showing a great deal. However, in a subgroup of individuals, a biological vulnerability may result in impairment of control that leads to behavioral excess or disinhibition and is associated with significant levels of distress and functional impairment. Consideration that shopping is universal and making an unwise purchase from time to time is common, although research has shown that there are individuals whose compulsive buying is extreme and leads to significant distress and impairment. Using scores on the Compulsive Buying Scale (1) of 2 standard deviations below the mean, Koran et al. estimated the prevalence of compulsive buying to be 5.8%; even with a very strict criterion of 3 standard deviations below the mean, the prevalence would be 1.4%. Previous estimates based on smaller, less representative samples have ranged from 1.8% to 16%. Thus, whatever estimate is used, the prevalence is higher than or similar to disorders that receive considerable research and clinical attention, and it represents a sizable group suffering distress and or functional impairment. The impairment criteria are important because it is how compulsive buying as a disorder is differentiated from more normal, if excessive, buying. Koran et al. found that when using the criterion of 2 standard deviations on the Compulsive Buying Scale, the individuals had significantly more maladaptive shopping and buying attitudes and behaviors and more financial problems than the other respondents. The data for the group with 3 standard deviations shows consequences that were even more extreme. This sort of distribution applies to many disorders. As mentioned above, ADHD and social anxiety disorder are two examples. One might also look at a long-accepted disorder: major depressive disorder. Many people suffer from occasional sadness and days on which they are "blue," but that does not diminish the importance of recognizing, researching, and treating major depressive disorder.

One can ask if people are morally responsible for their behavior if they commit unethical acts because of what has been classified as a mental disorder? Similarly, if an individual diagnosed with an impulse control disorder does something illegal, is he or she responsible? Having a diagnosable disorder does not eliminate the moral or legal consequences of bad behavior, although courts can require that the individuals receive treatment in order to prevent a recurrence of the problem. This can be seen with alcoholism, which has long been considered a disorder. If an alcoholic has an accident while driving under the influence, that is not considered a mitigating circumstance but the courts can require that the individual undergo treatment for their alcohol problem, along with any other sentencing requirements. Viewing compulsive buying from a medical perspective and as a diagnosable mental disorder has several advantages. It might facilitate routine screening for the condition by mental health professionals, and perhaps, even inclusion of the disorder in national prevalence surveys, which would help define the true prevalence of the disorder. It might also lead to the study of vulnerability factors for the development of the disorder, better characterization of brain-based circuits, and the development of effective psychosocial and medication treatments. Although prevention of overdiagnosis or possible misuse of diagnostic labels is

important, these concerns should be balanced against the advancement of knowledge that could potentially lead to new treatments or prevention strategies for serious human problems.

Reference

1. Faber RJ, O'Guinn TC: A clinical screener for compulsive buying. *J Consumer Res* 1992; 19:459-469

**ERIC HOLLANDER, M.D.
ANDREA ALLEN, PH.D.**

Address correspondence and reprint requests to Dr. Hollander, Department of Psychiatry, Mt. Sinai School of Medicine, One Gustave L Levy Place, New York, NY 10029; eric.hollander@mssm.edu (e-mail.)

Dr. Hollander has been a consultant to Ortho-McNeil, Abbott, and Forest; and has received research grants from NIMH, NIDA, NINDS, and OPD-FDA. Dr. Freedman has reviewed this editorial and found no evidence of influence from these relationships. Dr. Allen reports no competing interests.

COVER LETTER

poster submission

Arousal, Executive Control and Decision Making in Compulsive Buying Disorder

Thomas Zoëga Ramsøy ^{1, 2 *}

Farah Qureshi Zuraigat ¹

Catrine Jacobsen ¹

Dalia Bagdžiūnaitė ¹

Maiken Klindt Christensen ¹

Martin Skov ¹

Antoine Bechara ⁴

1. Decision Neuroscience Research Group, Department of Marketing, Copenhagen Business School, Frederiksberg, Denmark

2. Danish Research Centre for Magnetic Resonance, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark

3. DTU Informatics, Technical University of Denmark, Lyngby, Denmark

4. Department of Psychology, University of Southern California, Loa Angeles, California, USA

* Corresponding author: TZ Ramsøy, DNRG, Department of Marketing, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark. E-mail: tzr.marktg@cbs.dk; phone: (+45) 2181 1945

ABSTRACT

Compulsive buying disorder (CBD) is noted by an obsession with shopping and a chronic, repetitive purchasing behavior with adverse consequences for the sufferer and their social surroundings. While CBD is often classified as an impulse control disorder (ICD), little is still known about the actual psychological and physiological mechanisms underlying the phenomenon. By comparing subjects with CBD to a control group, we find that CBD is not associated with lower performance on executive function or emotional responses. Rather, an observed increase in willingness to pay (WTP) specifically for fashion products was associated with a stronger emotional response in CBD subjects, while no relationship between emotions and WTP could be observed in healthy controls. This suggests that CBD, instead of being tentatively classified as an ICD, should rather be understood as a behavioral addiction. By this token, products of interest (e.g. fashion items) produce bottom-up emotional responses that skews the decision-making process, leading CBD sufferers to make bad purchase decisions.

Arousal, Executive Control and Decision Making in Compulsive Buying Disorder

EXTENDED ABSTRACT

Compulsive buying disorder (CBD) is a condition characterized by an obsession with shopping and a chronic, repetitive purchasing behavior that has adverse consequences both for the sufferer and their social surroundings. Today, the prevalence of CBD is estimated to be somewhere between 5% and 7% (Koran, Faber, Aboujaoude, Large, & Serpe, 2006; Mueller et al., 2010b), and it is often reported that CBD is significantly more prevalent in women than men, although other studies report relatively equal distributions (, 2010b). While CBD is not formally classified as a psychiatric disorder in diagnostic manuals, it shares characteristics of other clinical disorders such as pathological gambling, including impulsivity, dysexecutive functions and mood disorders (Di Nicola et al., 2010; de Ruiter et al., 2009; van Holst, van den Brink, Veltman, & Goudriaan, 2010). As with pathological gambling (Kessler et al., 2008), CBD has also been reported to be characterized by comorbidities including Axis I disorders such as mood and especially anxiety disorders (Mueller et al., 2010a), lending further support to a possible common foundation between pathological gambling and CBD.

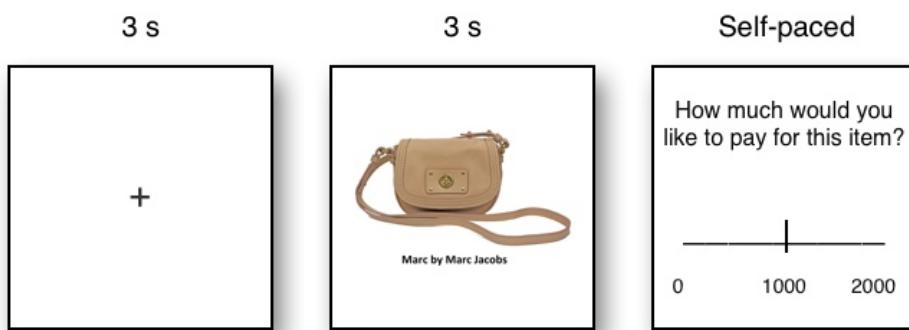
However, little is known about the neurobiological and neurophysiological underpinnings of CBD (Lejoyeux & Weinstein, 2010). On the one hand, studies have suggested that CBD is related to lower executive control and self-control (Claes et al., 2010), while others have reported emotion-related issues, such as elevated levels of anxiety, depression and other mood disorders (Black, 2001; Mueller et al., 2010a, 2010b). Thus, it is still unknown whether CBD is caused by a failing executive system, altered emotion responses, or both. Moreover, advances in understanding drug addiction have suggested that addiction behaviors could be the result of a “failing willpower” (Bechara, 2005), in which sufferers may display an aberrant relationship between emotional responses and value-based decision making. Taken together, these suggest at least three different, yet not mutually exclusive, causal mechanisms underlying CBD.

The aim of the present study was to provide a better insight into the basic neuropsychological and neurophysiological foundations of CBD. Based on the prior literature, three specific hypotheses were formed. The hypotheses pertained to 1) executive control, 2) emotional responses per se, and 3) the effect of emotional on decision making. Following the literature on impulse control disorders (Black, Shaw, & Blum, 2010; Lejoyeux & Weinstein, 2010; Regard, Knoch, Gütling, & Landis, 2003; van Holst et al., 2010), CBD could be related to dysexecutive function and a general lack of impulse control, and thus we would expect that CBD sufferers would score lower on tests of executive control, when compared to healthy subjects.

Sixty-three women (age range 19-51, mean/std = 26.5±7.5, all right handed) were recruited from the Copenhagen and Frederiksberg regions, Denmark, using both online (www.forsoegsperson.dk and www.videnskab.dk) and direct recruitment procedures. At enrollment, subjects read and signed an informed consent, and were initially informed and trained with the experimental procedure. They also filled out questionnaires relating to their overall health state, including the Compulsive Buying Scale (CBS) (Faber & O'Guinn, 1992a).

All stimuli were presented on a screen running with a 1920x1200 pixel resolution. Subjects were placed at an approximate distance of 60 cm from the screen, and underwent a

fixation calibration. The test consisted of three phases: i) an ISI phase which displayed a fixation cross for 3 seconds; ii) product image phase in which a product from one of four categories (purses, clothes, women's shoes, and fast-moving consumer goods) were shown for three seconds; and iii) a rating phase, in which subjects were asked to report how much they would like to pay for the product they had just seen, by using an on-screen visual analogue scale ranging from zero to 2.000 Danish Kroner ($\approx \$330$). The experimental design is illustrated in Figure 1. To increase the external validity of the test, subjects were instructed that the choices from two of the subjects from the cohort would be randomly selected and given 1.500 DKK each, and that five of each subject's choices would be randomly selected, and the product receiving the highest bid of those five would be realized. Should the highest bid not amount to 1.500 DKK, they would be paid the remaining amount in cash. This meant that subjects were motivated to optimize their product choices, which allowed us to better estimate the actual WTP, instead of subjective estimates of WTP.

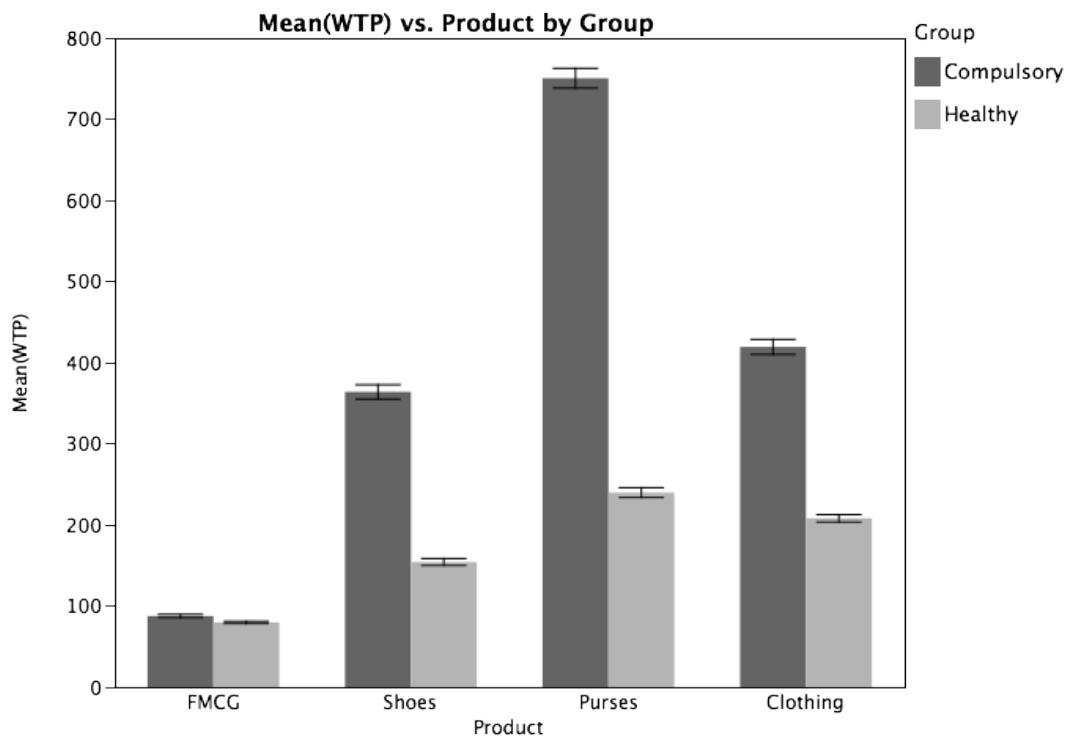


CAPTION: The shopping task. Subjects first saw a fixation cross, followed by the display of a product, and then a self-paced WTP task. Here, subjects decided how much they were willing to actually pay for the product, ranging from zero to DKK 2.000, by moving a visual analogue slider on the screen

Subjects were grouped according to their CBS score into three groups: healthy (range: 0-32), compensatory (range: 33-42), compulsive (range: 43-54) (Faber & O'Guinn, 1992b). The groups thus differed significantly on the CBS (one-way ANOVA $F=147.2$, $p<0.0001$) where the healthy individuals had the lowest CBS score ($n=23$, $mean \pm std = 26.1 \pm 3.6$), followed by the compensatory group ($n=19$, $mean \pm std = 37.2 \pm 2.6$) and finally the compulsive group ($n=9$, $mean \pm std = 46.8 \pm 3.4$). In the following analyses, we focus on direct comparisons between the CBD group and healthy controls.

Willingness To Pay

We first analyzed whether compulsive buyers showed an altered WTP for products compared to healthy controls by running an independent samples t-test. Here, we find a significant group effect, in that the CBD group displayed a disproportionately higher WTP ($mean \pm SEM = 402.22 \pm 1.99$ DKK) than the control group (171.94 ± 1.4 , independent sample t-test = -81.6, $p<0.0001$). Further post-hoc testing showed that this difference was significant for certain products, such as shoes, clothing and especially purses, but not for FMCG (see Figure 2).



Executive functions

Second, to compare whether CBD was related to dysfunctional executive processes, we analyzed the relationship between CBD and test results. Here, the neuropsychological tests were first analyzed using direct comparison between the CBD group and the healthy volunteers, using an independent samples t-test. In doing so, we find no group differences on the executive tests (Table 1), and only the Stroop response time (RT) shows a significant group effect, as the CBD group had significantly faster RT than the control group.

TEST	T	KS	p
Eriksen Flanker Test		0.12	0.731
Stroop Test			
- normal	-2.59		0.0078*
- interference	-1.33		0.0965(*)
- difference	0.28		0.389
Visual reaction		0.17	0.348
Go/No-go			
- result		0.12	0.771

TEST	T	KS	p
- time		0.17	0.289
- combined		0.12	0.71

* = significance at p<0.05 level; (*)= trend significance at p<0.1

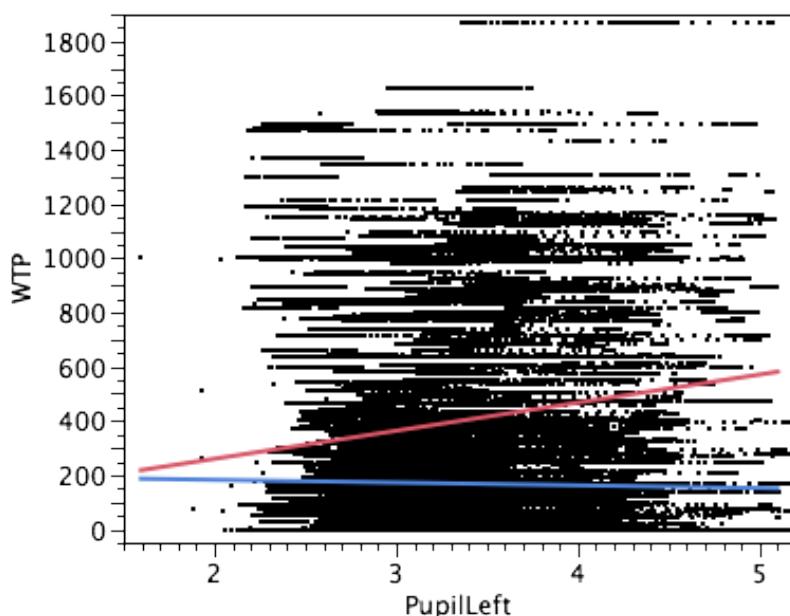
Emotional responses

We then turned to the effect of CBD on emotional responses. Here, we analyzed the relationship between the CBD and emotional arousal during product viewing by comparing emotional responses in the CBD group to the healthy subjects using an independent samples t-test. CBD subjects had a lower pupil dilation responses (3.23 ± 0.01) than controls (3.26 ± 0.01 , independent samples t-test=6.03, p<0.0001).

From emotion to decision

Finally, we wanted to test whether compulsive buying was associated with an alteration in the impact that emotions could have on WTP choice. Here, we ran a linear regression analysis using WTP as the dependent variable and using pupil dilation, group and the interaction between pupil dilation and group as the independent variables.

Here, we find that the effects of pupil dilation on WTP was significantly different between the groups. While controls show little relationship between pupil dilation and WTP, CBD subjects showed a strong relationship (Figure 3).



CAPTION: Relationship between pupil dilation response and WTP for CBD (red) and controls (blue)

Discussion

These results provide novel and compelling insights into the mechanisms of compulsive buying disorder. Notably, our data suggest that CBD is not due to a difference in executive functions, but rather a specifically stronger effect of arousal in making WTP choices. Notably, CBD subjects did not show a general stronger emotional response, but still showed a stronger effect of such responses on WTP choices.

Together, these results suggest that CBD may be erroneously linked to impulse control disorders. Rather, it seems that CBD is due to a stronger influence of emotions on decision-making. One possibility is that CBD should rather be classified as a behavioral addiction, in which urges and impulses have a stronger impact on decisions (Bechara, 2005).

References

- Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience*, 8(11), 1458-63. doi:10.1038/nn1584
- Black, D. W. (2001). Compulsive buying disorder: Definition, assessment, epidemiology and clinical management. *CNS Drugs*, 15(1), 17-27.
- Black, D. W., Shaw, M., & Blum, N. (2010). Pathological gambling and compulsive buying: Do they fall within an obsessive-compulsive spectrum? *Dialogues in Clinical Neuroscience*, 12(2), 175-85.
- Claes, L., Bijttebier, P., Eynde, F. V. D., Mitchell, J. E., Faber, R., Zwaan, M. D., & Mueller, A. (2010). Emotional reactivity and self-regulation in relation to compulsive buying. *Personality and Individual Differences*, 49(5), 526-530. doi:10.1016/j.paid.2010.05.020
- Di Nicola, M., Tedeschi, D., Mazza, M., Martinotti, G., Harnic, D., Catalano, V., . . . Janiri, L. (2010). Behavioural addictions in bipolar disorder patients: Role of impulsivity and personality dimensions. *Journal of Affective Disorders*, 125(1-3), 82-8. doi:10.1016/j.jad.2009.12.016
- Faber, R. J., & O'Guinn, T. C. (1992a). A clinical screener for compulsive buying. *Journal of Consumer Research*, 459-469. Retrieved from Google Scholar.
- Faber, R. J., & O'Guinn, T. C. (1992b). A clinical screener for compulsive buying. *Journal of Consumer Research*, 19(3), 459-69. Retrieved from <http://ideas.repec.org/a/ucp/jconrs/v19y1992i3p459-69.html>
- Kessler, R. C., Hwang, I., LaBrie, R., Petukhova, M., Sampson, N. A., Winters, K. C., & Shaffer, H. J. (2008). DSM-IV pathological gambling in the national comorbidity survey replication. *Psychological Medicine*, 38(9), 1351-60. doi:10.1017/S0033291708002900
- Koran, L., Faber, R., Aboujaoude, E., Large, M., & Serpe, R. (2006). Estimated prevalence of compulsive buying behavior in the united states. *American Journal of Psychiatry*, 163(10), 1806-1812. Retrieved from Google Scholar.
- Lejoyeux, M., & Weinstein, A. (2010). Compulsive buying. *The American Journal of Drug and Alcohol Abuse*, 36(5), 248-53. doi:10.3109/00952990.2010.493590
- Mueller, A., Mitchell, J. E., Black, D. W., Crosby, R. D., Berg, K., & de Zwaan, M. (2010a). Latent profile analysis and comorbidity in a sample of individuals with compulsive buying disorder. *Psychiatry Research*, 178(2), 348-53. doi:10.1016/j.psychres.2010.04.021
- Mueller, A., Mitchell, J. E., Crosby, R. D., Gefeller, O., Faber, R. J., Martin, A., . . . de Zwaan, M. (2010b). Estimated prevalence of compulsive buying in germany and its association with sociodemographic characteristics and depressive symptoms. *Psychiatry Research*, 180(2-3), 137-42. doi:10.1016/j.psychres.2009.12.001
- Regard, M., Knoch, D., Gütling, E., & Landis, T. (2003). Brain damage and addictive behavior: A neuropsychological and electroencephalogram investigation with pathologic gamblers. *Cogn Behav Neurol*, 16(1), 47-53.

de Ruiter, M. B., Veltman, D. J., Goudriaan, A. E., Oosterlaan, J., Sjoerds, Z., & van den Brink, W. (2009). Response perseveration and ventral prefrontal sensitivity to reward and punishment in male problem gamblers and smokers. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, 34(4), 1027-38. doi: 10.1038/npp.2008.175

van Holst, R. J., van den Brink, W., Veltman, D. J., & Goudriaan, A. E. (2010). Why gamblers fail to win: A review of cognitive and neuroimaging findings in pathological gambling. *Neuroscience and Biobehavioral Reviews*, 34(1), 87-107. doi:10.1016/j.neubiorev.2009.07.007

A review of compulsive buying disorder

DONALD W. BLACK

Department of Psychiatry, University of Iowa Roy J. and Lucille A. Carver College of Medicine, Iowa City, IA 52242, USA

Compulsive buying disorder (CBD) is characterized by excessive shopping cognitions and buying behavior that leads to distress or impairment. Found worldwide, the disorder has a lifetime prevalence of 5.8% in the US general population. Most subjects studied clinically are women (~80%), though this gender difference may be artifactual. Subjects with CBD report a preoccupation with shopping, pre-purchase tension or anxiety, and a sense of relief following the purchase. CBD is associated with significant psychiatric comorbidity, particularly mood and anxiety disorders, substance use disorders, eating disorders, and other disorders of impulse control. The majority of persons with CBD appear to meet criteria for an Axis II disorder, although there is no special "shopping" personality. Compulsive shopping tends to run in families, and these families are filled with mood and substance use disorders. There are no standard treatments. Psychopharmacologic treatment studies are being actively pursued, and group cognitive-behavioral models have been developed and are promising. Debtors Anonymous, simplicity circles, bibliotherapy, financial counseling, and marital therapy may also play a role in the management of CBD.

Key words: Compulsive shopping, compulsive buying, impulse control disorders

(World Psychiatry 2007;6:14-18)

Compulsive buying disorder (CBD) was first described clinically in the early 20th century by Bleuler (1) and Kraepelin (2), both of whom included CBD in their textbooks. Bleuler writes: "As a last category Kraepelin mentions the buying maniacs (oniomaniacs) in whom even buying is compulsive and leads to senseless contraction of debts with continuous delay of payment until a catastrophe clears the situation a little – a little bit never altogether because they never admit to their debts" (1). Bleuler described CBD as an example of a "reactive impulse", or "impulsive insanity", which he grouped alongside kleptomania and pyromania.

CBD attracted little attention throughout the 20th century except among consumer behaviorists (3-6) and psychoanalysts (7-9). Interest was revived in the early 1990s, when clinical case series from three independent research groups appeared (10-12). The disorder has been described worldwide, with reports coming from the US (10-12), Canada (5), England (4), Germany (6), France (13), and Brazil (14).

The appropriate classification of CBD continues to be debated. Some researchers have linked CBD to addictive disorders (15), while others have linked it to obsessive-compulsive disorder (16), and still others to mood disorders (17). While not included in DSM-IV (18), CBD was included in DSM-III-R (19) as an example of an "impulse-control disorder not otherwise specified". Research criteria have been developed that emphasize its cognitive and behavioral aspects (10). Some writers have criticized attempts to categorize CBD as an illness, which they see as part of a trend to "medicalize" behavioral problems (20). Yet, this approach ignores the reality of CBD, and both trivializes and stigmatizes attempts to understand or treat the disorder.

EPIDEMIOLOGY

Koran et al (21) recently estimated the point prevalence

of CBD to be 5.8% of respondents, based on results from a random telephone survey of 2,513 adults conducted in the US. Earlier, Faber and O'Guinn (22) had estimated the prevalence of CBD to fall between 2% and 8% of the general population of Illinois. Both research groups had used the Compulsive Buying Scale (CBS) (23) to identify compulsive buyers. Other surveys have reported figures ranging from 12% to 16% (24,25). There is no evidence that CBD has increased in prevalence in the past few decades.

Community based and clinical surveys suggest that 80% to 95% of persons with CBD are women (10-12,23). The reported gender difference could be artifactual: women readily acknowledge that they enjoy shopping, whereas men are more likely to report that they "collect". The report of Koran et al (21) suggests that this may be the case: in their survey, a near equal percentage of men and women met criteria for CBD (5.5% and 6.0%, respectively). However, Dittmar (26) concluded from a general population survey in the United Kingdom, in which 92% of respondents considered compulsive shoppers were women, that the gender difference is real and is not an artifact of men being underrepresented in samples.

The age of onset of CBD appears to be in the late teens or early twenties (11,12,27), though McElroy et al (10) reported a mean age at onset of 30 years. It may be that the age of onset corresponds with emancipation from the home, and the age at which people first establish credit accounts.

There are no careful longitudinal studies of CBD, but the majority of subjects studied by Schlosser et al (12) and McElroy et al (10) describe their course as continuous. Aboujaoude et al (28) suggested that persons with CBD who responded to treatment with citalopram were likely to remain in remission during one-year follow-up, a finding that suggests that treatment could alter the natural history of the disorder. The authors' personal observation is that subjects with CBD typically report decades of compulsive shopping

behavior at the time of presentation, although it might be argued that clinical samples are biased in favor of severity.

There is some evidence that CBD runs in families and that within these families mood, anxiety, and substance use disorders are excessive. McElroy et al (8) reported that, of 18 individuals with CBD, 17 had one or more first-degree relatives (FDRs) with major depression, 11 with an alcohol or drug use disorder, and three with an anxiety disorder. Three had relatives with CBD. Black et al (29) used the family history method to assess 137 FDRs of 33 persons with CBD. FDRs were significantly more likely than those in a comparison group to have depression, alcoholism, a drug use disorder, "any" psychiatric disorder, and "more than one psychiatric disorder". CBD was identified in 9.5% of the FDRs of the CBD probands (CBD was not assessed in the comparison group). In molecular genetic studies, Devor et al (30) failed to find an association between two serotonin transporter gene polymorphisms and CBD, while Comings (31) reported an association of CBD with the DRD1 receptor gene.

CLINICAL SYMPTOMS

Persons with CBD are preoccupied with shopping and spending, and devote significant time to these behaviors. While it might be argued that a person could be a compulsive shopper and not spend, and confine his or her interest to window shopping, this pattern is uncommon. The author's personal observation is that the two aspects – shopping and spending – are intertwined. Persons with CBD often describe an increasing level of urge or anxiety that can only lead to a sense of completion when a purchase is made.

The author has been able to identify four distinct phases of CBD: 1) anticipation; 2) preparation; 3) shopping; and 4) spending. In the first phase, the person with CBD develops thoughts, urges, or preoccupations with either having a specific item, or with the act of shopping. In the second phase, the person prepares for shopping and spending. This can include decisions on when and where to go, on how to dress, and even which credit cards to use. Considerable research may have taken place about sale items, new fashions, or new shops. The third phase involves the actual shopping experience, which many individuals with CBD describe as intensely exciting, and can even lead to a sexual feeling (12). Finally, the act is completed with a purchase, often followed by a sense of let down, or disappointment with oneself (21). In a study of the antecedents and consequences of CBD, Miltenberger et al (32) reported that negative emotions (e.g., depression, anxiety, boredom, self-critical thoughts, anger) were the most commonly cited antecedents to CBD, while euphoria or relief from the negative emotions were the most common consequence.

Individuals with CBD tend to shop by themselves, although some will shop with friends who may share their interest in shopping (11,12). In general, CBD is a private pleasure which could lead to embarrassment if someone

not similarly interested in shopping accompanied them. Shopping may occur in just about any venue, ranging from high fashion department stores and boutiques to consignment shops or garage sales. Income has relatively little to do with the existence of CBD: persons with a low income can still be fully preoccupied by shopping and spending, although their level of income will lead them to shop at a consignment shop rather than a department store.

Typical items purchased by persons with CBD include (in descending order) clothing, shoes, compact discs, jewelry, cosmetics, and household items (11,12,32). Individually, the items purchased by compulsive shoppers tend not to be particularly expensive, but the author has observed that many compulsive shoppers buy in quantity resulting in out of control spending. Anecdotally, patients often report buying a product based on its attractiveness or because it was a bargain. In the study by Christenson et al (11), compulsive shoppers reported spending an average of \$110 during a typical shopping episode compared with \$92 reported in the study by Schlosser et al (12).

Although research has not identified gender specific buying patterns, in the author's experience men tend to have a greater interest than women in electronic, automotive, or hardware goods. Like women, they are also interested in clothing, shoes, and compact discs.

Subjects generally are willing to acknowledge that CBD is problematic. Schlosser et al (10) reported that 85% of their subjects expressed concern with their CBD-related debts, and that 74% felt out of control while shopping. In the study by Miltenberger et al (32), 68% of persons with CBD reported that it negatively affected their relationships. Christenson et al (11) reported that nearly all of their subjects (92%) tried to resist their urges to buy, but were rarely successful. The subjects indicated that 74% of the time they experienced an urge to buy, the urge resulted in a purchase.

CBD tends to occur year round, although it may be more problematic during the Christmas or other important holidays, and around the birthdays of family members and friends (12). Schlosser et al (12) found that subjects reported a range of behaviors regarding the outcome of a purchase, including returning the item, failing to remove the item from the packaging, selling the item, or even giving it away.

In a study of 44 subjects with CBD, Black et al (33) reported that greater severity was associated with lower gross income, less likelihood of having an income above the median, and spending a lower percentage of income on sale items. Subjects with more severe CBD were also more likely to have comorbid Axis I or Axis II disorders. These data suggest that the most severe forms of CBD are found in persons with low incomes who have little ability to control or to delay their urge to make impulsive purchases.

PSYCHIATRIC COMORBIDITY

Persons with CBD frequently meet criteria for Axis I dis-

orders, particularly mood disorders (21-100%) (27,34), anxiety disorders (41-80%) (10,12), substance use disorders (21-46%) (11,29), and eating disorders (8-35%) (10,27). Disorders of impulse control are also relatively common in these individuals (21-40%) (10,11).

Schlosser et al (12) found that nearly 60% of subjects with CBD met criteria for at least one Axis II disorder. While there was no special "shopping" personality, the most frequently identified personality disorders were the obsessive-compulsive (22%), avoidant (15%), and borderline (15%) types. Krueger (7), a psychoanalyst, described four patients who he observed to have aspects of narcissistic character pathology.

Etiology

The etiology of CBD is unknown, though speculation has settled on developmental, neurobiological, and cultural influences. Psychoanalysts (7-9) have suggested that early life events, such as sexual abuse, are causative factors. Yet, no special or unique family constellation or pattern of early life events has been identified in persons with CBD.

Neurobiological theories have centered on disturbed neurotransmission, particularly involving the serotonergic, dopaminergic, or opioid systems. Selective serotonin reuptake inhibitors (SSRIs) have been used to treat CBD (27,34-38), in part because investigators have noted similarities between CBD and obsessive-compulsive disorder, a disorder known to respond to SSRIs. Dopamine has been theorized to play a role in "reward dependence", which has been claimed to foster "behavioral addictions" (e.g., CBD, pathological gambling) (39). Case reports suggesting benefit from the opiate antagonist naltrexone have led to speculation about the role of opiate receptors (40,41). There is currently no direct evidence to support the role of these neurotransmitter systems in the etiology of CBD.

Cultural mechanisms have been proposed to recognize the fact that CBD occurs mainly in developed countries (42). Elements which appear necessary for the development of CBD include the presence of a market-based economy, the availability of a wide variety of goods, disposable income, and significant leisure time. For these reasons, CBD is unlikely to occur in poorly developed countries, except among the wealthy elite (Imelda Marcos and her many shoes come to mind).

Assessment

The goal of assessment is to identify CBD through inquiries regarding the person's attitudes and behaviors towards shopping and spending (43). Inquiries might include: "Do you feel overly preoccupied with shopping and spending?"; "Do you ever feel that your shopping behavior is excessive, inappropriate or uncontrolled?"; "Have your shopping desires, urges, fantasies, or behaviors ever been overly time consuming, caused you to feel upset or guilty, or lead

to serious problems in your life such as financial or legal problems or the loss of a relationship?".

Clinicians should note past psychiatric treatment, including medications, hospitalizations, and psychotherapy. A history of physical illness, surgical procedures, drug allergies, or medical treatment is important to note, because it may help rule out medical explanations as a cause of the CBD (e.g., neurological disorders, brain tumors). Bipolar disorder needs to be ruled out as a cause of the excessive shopping and spending. Typically, the manic patient's unrestrained spending corresponds to manic episodes, and is accompanied by euphoric mood, grandiosity, unrealistic plans, and often a giddy, expansive affect. The pattern of shopping and spending in the person with CBD lacks the periodicity seen with bipolar patients, and suggests an ongoing preoccupation.

Normal buying behavior should also be ruled out. In the US and other developed countries, shopping is a major pastime, particularly for women, and frequent shopping does not necessarily constitute evidence in support of a diagnosis of CBD. Normal buying can sometimes take on a compulsive quality, particularly around special holidays or birthdays. Persons who receive an inheritance or win a lottery may experience shopping sprees as well.

Several instruments have been developed to either identify CBD or rate its severity. The CBS (23), already mentioned, consists of seven items representing specific behaviors, motivations, and feelings associated with compulsive buying, and reliably distinguishes normal buyers from those with CBD. Edwards (44) has developed a useful 13-item scale that assesses important experiences and feelings about shopping and spending. Monahan et al (45) modified the Yale Brown Obsessive-Compulsive Scale to create the YBOCS-Shopping Version (YBOCS-SV) to assess cognitions and behaviors associated with CBD. This 10-item scale rates time involved, interference, distress, resistance, and degree of control for both cognitions and behaviors. The instrument is designed to measure severity of CBD, and change during clinical trials.

Treatment

There are no evidence-based treatments for CBD. In recent years, treatment studies of CBD have focused on the use of psychotropic medication (mainly antidepressants) and cognitive-behavioral therapy (CBT).

Interest in CBT has largely replaced earlier interest in psychodynamic therapies. Several competing CBT models have been developed, the most successful involving the use of group treatment (46-49). The first use of group therapy was described by Damon (46). Subsequent group models were developed by Burgard and Mitchell (47), Villarino et al (48), and more recently by Benson and Gengler (49). Mitchell et al (50) reported that their group CBT model produced significant improvement compared to a wait list in a 12-week pilot study; improvement was maintained during a 6-months follow-up. Benson (51) has recently developed a

comprehensive self-help program which combines cognitive-behavioral strategies with self-monitoring. A detailed workbook, a shopping diary, and a CD-ROM are included.

Several self-help books (bibliotherapy) are available (52-54), and may be helpful to some persons with CBD. Debtors Anonymous, patterned after Alcoholics Anonymous, is a voluntary, lay-run group that provides an atmosphere of mutual support and encouragement for those with substantial debts. Simplicity circles are available in some US cities; these voluntary groups encourage people to adopt a simple lifestyle, and to abandon their CBD (55). Many subjects with CBD develop substantial financial problems, and may benefit from financial counseling (56). The author has seen cases in which a financial conservator has been appointed to control the patient's finances, and appears to have helped. While a conservator controls the person's spending, this approach does not reverse his or her preoccupation with shopping and spending. Marriage (or couples) counseling may be helpful, particularly when CBD in one member of the dyad has disrupted the relationship (57).

Psychopharmacologic treatment studies have yielded mixed results. An early case series suggested that antidepressants could curb CBD (58), and an early open-label trial using fluvoxamine showed benefit (34). Yet, two subsequent randomized controlled trials found that fluvoxamine did no better than placebo (35,36). In another open-label trial (28), citalopram produced substantial improvement. In this particular study, responders to open-label citalopram were then enrolled in a nine-week randomized placebo controlled trial (38). Compulsive shopping symptoms returned in five of eight subjects assigned to placebo compared with none of the seven who continued taking citalopram. By comparison, escitalopram showed little effect for CBD in an identically designed discontinuation trial by the same investigators (39). Grant (40) and Kim (41) have described cases in which persons with CBD improved with naltrexone, suggesting that opiate antagonists might play a role in the treatment of CBD. Interpretation of treatment studies is complicated by the high placebo response rate associated with CBD (ranging to 64%) (35).

The author has developed a set of recommendations (59). First, pharmacologic treatment trials provide little guidance, and patients should be informed that they cannot rely on medication. Further, patients should: a) admit that they have CBD; b) get rid of credit cards and checkbooks, because they are easy sources of funds that fuel the disorder; c) shop with a friend or relative; the presence of a person without CBD will help curb the tendency to overspend; and d) find meaningful ways to spend one's leisure time other than shopping.

References

- Bleuler E. Textbook of psychiatry. New York: Macmillan, 1930.
- Kraepelin E. Psychiatrie, 8th ed. Leipzig: Barth, 1915.
- O'Guinn TC, Faber RJ. Compulsive buying: a phenomenological exploration. *J Consumer Res* 1989;16:147-57.
- Elliott R. Addictive consumption: function and fragmentation in post-modernity. *J Consumer Policy* 1994;17:159-79.
- Valence G, D'Astous A, Fortier L. Compulsive buying: concept and measurement. *J Consum Policy* 1988;11:419-33.
- Scherhorn G, Reisch LA, Raab G. Addictive buying in West Germany: an empirical study. *J Consum Policy* 1990;13:355-87.
- Krueger DW. On compulsive shopping and spending: a psychodynamic inquiry. *Am J Psychother* 1988;42:574-84.
- Lawrence L. The psychodynamics of the compulsive female shopper. *Am J Psychoanal* 1990;50:67-70.
- Winestine MC. Compulsive shopping as a derivative of childhood seduction. *Psychoanal Q* 1985;54:70-2.
- McElroy S, Keck PE Jr, Pope HG Jr et al. Compulsive buying: a report of 20 cases. *J Clin Psychiatry* 1994;55:242-8.
- Christenson GA, Faber JR, de Zwann M. Compulsive buying: descriptive characteristics and psychiatric comorbidity. *J Clin Psychiatry* 1994;55:5-11.
- Schlosser S, Black DW, Repertinger S et al. Compulsive buying: demography, phenomenology, and comorbidity in 46 subjects. *Gen Hosp Psychiatry* 1994;16:205-12.
- Lejoyeux M, Tassain V, Solomon J et al. Study of compulsive buying in depressed patients. *J Clin Psychiatry* 1997;58:169-73.
- Bernik MA, Akerman D, Amaral JAMS et al. Cue exposure in compulsive buying. *J Clin Psychiatry* 1996;57:90.
- Krych R. Abnormal consumer behavior: a model of addictive behaviors. *Adv Consum Res* 1989;16:745-8.
- Hollander E (ed). Obsessive-compulsive related disorders. Washington: American Psychiatric Press, 1993.
- Lejoyeux M, Andes J, Tassain V et al. Phenomenology and psychopathology of uncontrolled buying. *Am J Psychiatry* 1996;152:1524-9.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed, revised. Washington: American Psychiatric Press, 1987.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed, text revision. Washington: American Psychiatric Publishing, 2004.
- Lee S, Mysyk A. The medicalization of compulsive buying. *Soc Sci Med* 2004;58:1709-18.
- Koran LM, Faber RJ, Aboujaoude E et al. Estimated prevalence of compulsive buying in the United States. *Am J Psychiatry* 2006;163:1806-12.
- Faber RJ, O'Guinn TC. Classifying compulsive consumers: advances in the development of a diagnostic tool. *Adv Consum Res* 1989;16:147-57.
- Faber RJ, O'Guinn TC. A clinical screener for compulsive buying. *J Consumer Res* 1992;19:459-69.
- Magee A. Compulsive buying tendency as a predictor of attitudes and perceptions. *Adv Consum Res* 1994;21:590-4.
- Hassay DN, Smith CL. Compulsive buying: an examination of consumption motive. *Psychol Marketing* 1996;13:741-52.
- Dittmar H. Understanding and diagnosing compulsive buying. In: Coombs R (ed). Addictive disorders: a practical handbook. New York: Wiley, 2004:411-50.
- Koran LM, Bullock KD, Hartston HJ et al. Citalopram treatment of compulsive shopping: an open-label study. *J Clin Psychiatry* 2002;63:704-8.
- Aboujaoude E, Gamel N, Koran LM. A 1-year naturalistic follow-up of patients with compulsive shopping disorder. *J Clin Psychiatry* 2003;64:946-50.
- Black DW, Repertinger S, Gaffney GR et al. Family history and psychiatric comorbidity in persons with compulsive buying: preliminary findings. *Am J Psychiatry* 1998;155:960-3.
- Devor EJ, Magee HJ, Dill-Devor RM et al. Serotonin transporter gene (5-HTT) polymorphisms and compulsive buying. *Am J Med Genet* 1999;88:123-5.

31. Comings DE. The molecular genetics of pathological gambling. *CNS Spectrums* 1998;6:20-37.
32. Miltenberger RG, Redlin J, Crosby R et al. Direct and retrospective assessment of factors contributing to compulsive buying. *J Behav Ther Exp Psychiatry* 2003;34:1-9.
33. Black DW, Monahan P, Schlosser S et al. Compulsive buying severity: an analysis of compulsive buying scale results in 44 subjects. *J Nerv Ment Dis* 2001;189:123-7.
34. Black DW, Monahan P, Gabel J. Fluvoxamine in the treatment of compulsive buying. *J Clin Psychiatry* 1997;58:159-63.
35. Black DW, Gabel J, Hansen J et al. A double-blind comparison of fluvoxamine versus placebo in the treatment of compulsive buying disorder. *Ann Clin Psychiatry* 2000;12:205-11.
36. Ninan PT, McElroy SL, Kane CP et al. Placebo-controlled study of fluvoxamine in the treatment of patients with compulsive buying. *J Clin Psychopharmacol* 2000;20:362-6.
37. Koran LM, Chuang HW, Bullock KD et al. Citalopram for compulsive shopping disorder: an open-label study followed by a double-blind discontinuation. *J Clin Psychiatry* 2003;64:793-8.
38. Koran LM. Escitalopram treatment evaluated in patients with compulsive shopping disorder. *Primary Psychiatry* 2005;12:13.
39. Holden C. Behavioral addictions; do they exist? *Science* 2001;294: 980-2.
40. Grant JE. Three cases of compulsive buying treated with naltrexone. *Int J Psychiatry Clin Pract* 2003;7:223-5.
41. Kim SW. Opioid antagonists in the treatment of impulse-control disorders. *J Clin Psychiatry* 1998;59:159-64.
42. Black DW. Compulsive buying disorder: definition, assessment, epidemiology and clinical management. *CNS Drugs* 2001;15:17-27.
43. Black DW. Assessment of compulsive buying. In: Benson A (ed). *I shop, therefore I am – compulsive buying and the search for self*. New York: Aronson, 2000:191-216.
44. Edwards EA. Development of a new scale to measure compulsive buying behavior. *Fin Counsel Plan* 1993;4:67-84.
45. Monahan P, Black DW, Gabel J. Reliability and validity of a scale to measure change in persons with compulsive buying. *Psychiatry Res* 1995;64:59-67.
46. Damon JE. *Shopaholics: serious help for addicted spenders*. Los Angeles: Price Stein Sloan, 1988.
47. Burgard M, Mitchell JE. Group cognitive-behavioral therapy for buying disorders. In: Benson A (ed). *I shop, therefore I am – compulsive buying and the search for self*. New York: Aronson, 2000:367-97.
48. Villarino R, Otero-Lopez JL, Casto R. *Adicion a la compra: análisis, evaluación y tratamiento*. Madrid: Ediciones Pirámide, 2001.
49. Benson A, Gengler M. Treating compulsive buying. In: Coombs R (ed). *Addictive disorders: a practical handbook*. New York: Wiley, 2004:451-91.
50. Mitchell JE, Burgard M, Faber R et al. Cognitive behavioral therapy for compulsive buying disorder. *Behav Res Ther* 2006;44:1859-65.
51. Benson A. Stopping overshopping – A comprehensive program to help eliminate overshopping. New York: April Benson, 2006; www.stoppingovershopping.com.
52. Arendon G. *Born to spend: how to overcome compulsive spending*. Blue Ridge Summit: Tab Books, 1991.
53. Catalano EM, Sonenberg N. *Consuming passions – help for compulsive shoppers*. Oakland: New Harbinger Publications, 1993.
54. Wesson C. *Women who shop too much: overcoming the urge to splurge*. New York: St. Martin's Press, 1991.
55. Andrews C. Simplicity circles and the compulsive shopper. In: Benson A (ed). *I shop, therefore I am – compulsive buying and the search for self*. New York: Aronson, 2000:484-96.
56. McCall K. Financial recovery counseling. In: Benson A (ed). *I shop, therefore I am – compulsive buying and the search for self*. New York: Aronson, 2000:457-83.
57. Mellan O. Overcoming overspending in couples. In: Benson A (ed). *I shop, therefore I am – compulsive buying and the search for self*. New York: Aronson, 2000:341-66.
58. McElroy S, Satlin A, Pope HG Jr et al. Treatment of compulsive shopping with antidepressants: a report of three cases. *Ann Clin Psychiatry* 1991;3:199-204.
59. Kuzma J, Black DW. Compulsive shopping – when spending begins to consume the consumer. *Current Psychiatry* 2006;7:27-40.

Pathological gambling and compulsive buying: do they fall within an obsessive-compulsive spectrum?

Donald W. Black, MD; Martha Shaw, BA; Nancee Blum, MSW



Both compulsive buying (CB) and pathological gambling (PG) have been proposed as members of a spectrum of disorders related to obsessive-compulsive disorder (OCD). The spectrum hypothesis originated in the early 1990s and has gained considerable support, despite the lack of empirical evidence. Interest in this hypothesis has become critical because some investigators have recommended the creation of a new category that includes these disorders in DSM-5, now under development. In this article, the authors describe the origin of the obsessive-compulsive (OC) spectrum and its theoretical underpinnings, review both CB and PG, and discuss the data both in support of and against an OC spectrum. Both disorders are described in terms of their history, definition, classification, phenomenology, family history, pathophysiology, and clinical management. The authors conclude that: (i) CB and PG are probably not related to OCD, and there is insufficient evidence to place them within an OC spectrum in DSM-V; (ii) PG should stay with the impulse-control disorders (ICDs); and (iii) a new diagnosis of CB should be created and be classified as an ICD.

© 2010, LLS SAS

Dialogues Clin Neurosci. 2010;12:175-185.

I

In the early 1990s, interest began to grow around the concept of an obsessive-compulsive (OC) spectrum. Hollander and others¹⁻³ wrote of a spectrum of disorders related to obsessive-compulsive disorder (OCD). Based on his experience as an OCD researcher, Hollander considered OCD to be at the center of the spectrum, and described its breadth and overlap with many other psychiatric disorders. These disorders were considered to lie along orthogonal axes of impulsivity vs compulsiveness, uncertainty vs certainty, and cognitive vs motoric (features). The OC spectrum concept was quickly embraced by other investigators because it offered a new way to think about the relationship among many neglected disorders, and it potentially offered new treatment options.^{4,5} Not all investigators have agreed, and several critical reviews have appeared.⁶⁻⁹

Despite the criticism, the concept of a group of disorders being related to OCD remains of great theoretical interest. The idea that disorders are related is crucial to classification schemes, and why should a group of disorders *not* be related to OCD? This question is now of singular interest because those responsible for developing the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) must decide whether to create a separate category for OCD and potentially related disorders, or to keep OCD with the anxiety disorders. If

Keywords: compulsive buying; pathological gambling; obsessive-compulsive spectrum; impulse control disorder; behavioral addiction

Author affiliations: Department of Psychiatry, University of Iowa Roy J. and Lucille A. Carver College of Medicine, Iowa City, Iowa, USA

Address for correspondence: Dr Donald Black, Psychiatry Research/2-126B MEB, University of Iowa Carver College of Medicine, Iowa City, Iowa 52242, USA
(e-mail: donald-black@uiowa.edu)

Clinical research

Selected abbreviations and acronyms

CB	<i>compulsive buying</i>
ICD	<i>impulse-control disorder</i>
OC	<i>obsessive-compulsive</i>
OCD	<i>obsessive-compulsive disorder</i>
PG	<i>pathological gambling</i>
SSRI	<i>selective serotonin reuptake inhibitor</i>

they create a new category for the OC spectrum they will need to determine its breadth.

The OC spectrum's boundaries have expanded or contracted according to the views of the investigator concerned. It has been described as including disorders of impulse control such as pathological gambling (PG), trichotillomania, and kleptomania; Tourette's and other tic disorders; impulsive personality disorders (eg, borderline personality disorder); hypochondriasis and body dysmorphic disorder; eating disorders; and several disorders not currently recognized in *DSM-IV-TR*¹⁰ such as compulsive buying (CB) and sexual addiction.¹⁴ Few investigators have offered evidence to validate a relationship among the disorders. Typically, such evidence might include comparisons of phenomenology, natural history, family history, biological markers, and treatment response.¹¹

OCD holds an important place at the center of the spectrum. Currently classified in *DSM-IV-TR*¹⁰ as an anxiety disorder, OCD is independent of other anxiety disorders in the International Classification of Diseases (ICD) system,¹² and a strong rationale has been presented by Zohar et al¹³ for its separation from these disorders. First, OCD often begins in childhood, whereas other anxiety disorders typically have a later age of onset. OCD has a nearly equal gender distribution, unlike the other anxiety disorders, which are more common in women. Studies of psychiatric comorbidity show that, unlike the other anxiety disorders, persons with OCD generally tend not to have elevated rates of substance misuse. Family studies have not shown a clear association between OCD and the other anxiety disorders. Brain circuitry that mediates OCD appears to be different from that involved in other anxiety disorders. Lastly, OCD is unique with regard to its response to the serotonin reuptake inhibitors (SSRIs), while noradrenergic medications, effective in mood disorders, and somewhat effective in anxiety disorders, are largely ineffective in OCD. On the other hand, the benzodiazepines, which have little effect on OCD, are often effective for the other anxiety disorders. Further, Zohar et al¹³ have argued that

recognizing the spectrum would contribute to improved classification, thus enabling a more precise description of endophenotype and biological markers that characterize these conditions, and that better classification could lead to more specific treatments.

Apart from the possibility of an OC spectrum, there has been no consistent approach to categorizing impulsive and compulsive disorders. While some have decried the "medicalization" of problematic behaviors such as CB,¹⁴ discussion has mainly focused on how these disorders should be classified, their relationship to other putative OC spectrum disorders, and whether some of them stand alone as independent disorders (eg, CB, compulsive sexual behavior).

Alternative classification schemes have emphasized the relationship of a putative OC spectrum disorder to depression or other mood disorders, to the impulse-control disorders (ICDs), or to the addictive disorders. Recently, it has been suggested that at least some of the disorders included in the OC spectrum be placed within a new diagnostic category that combines behavioral and substance addictions.¹⁵ "Behavioral addictions" include disorders that the National Institute on Drug Abuse (NIDA) considers to be relatively pure models of addiction because they are not contaminated by the presence of an exogenous substance.

With this background in mind, this article will focus on the status of PG and CB. Are these disorders part of an OC spectrum as defined by Hollander and coworkers? Are they more appropriately considered impulse control disorders (ICDs) or addictions? Are they related to one another? These and other questions will be considered as we explore CB, PG, and the OC spectrum.

Compulsive buying

CB has been described in the psychiatric nomenclature for nearly 100 years. German psychiatrist Emil Kraepelin¹⁶ wrote about the uncontrolled shopping and spending behavior called *oniomania* ("buying mania"). He was later quoted by Swiss psychiatrist Eugen Bleuler¹⁷ in his *Lehrbuch der Psychiatrie*:

As a last category, Kraepelin mentions the buying maniacs (*oniomaniacs*) in whom even buying is compulsive and leads to senseless contraction of debts with continuous delay of payment until a catastrophe clears the situation a little – a little bit never altogether because they never admit all their debts. The particular element is impul-

siveness; they cannot help it, which sometimes even expresses itself in the fact that notwithstanding a good school intelligence, the patients are absolutely incapable of thinking differently and conceiving the senseless consequences of their act, and the possibilities of not doing it." (p 540).

Kraepelin and Bleuler each considered "buying mania" an example of a *reactive impulse* or *impulsive insanity*, and placed it alongside kleptomania and pyromania. They may have been influenced by French psychiatrist Jean Esquirol's¹⁸ earlier concept of *monomania*, a term he used to describe otherwise normal persons who had some form of pathological preoccupation.

CB attracted little attention until the late 1980s and early 1990s when consumer behavior researchers showed the disorder to be widespread¹⁹⁻²¹ and descriptive studies appeared in the psychiatric literature.²²⁻²⁵ McElroy et al²² developed an operational definition that encompasses the cognitive and behavioral aspects of CB. Their definition requires evidence of impairment from marked subjective distress, interference in social or occupational functioning, or financial/legal problems. Further, the syndrome could not be attributed to mania or hypomania. Other definitions have come from consumer behavior researchers or social psychologists. Faber and O'Guinn²⁶ defined the disorder as "chronic buying episodes of a somewhat stereotyped fashion in which the consumer feels unable to stop or significantly moderate his behavior" (p 738). Edwards,²⁷ another consumer behaviorist, suggests that compulsive buying is an "abnormal form of shopping and spending in which the afflicted consumer has an overpowering uncontrollable, chronic and repetitive urge to shop and spend (that functions) ... as a means of alleviating negative feelings of stress and anxiety." (p 67). Dittmar²⁸ describes three cardinal features: irresistible impulse, loss of control, and carrying on despite adverse consequences. Some consumer behavior researchers consider CB part of a spectrum of aberrant consumer behavior, which includes pathological gambling, shoplifting, and credit abuse).²⁹

CB is not included in either the *DSM-IV-TR*¹⁰ or the World Health Organization *International Classification of Diseases, Tenth Edition*.¹² Whether to include CB in *DSM-5* is being debated.³⁰ McElroy et al²³ suggest that compulsive shopping behavior might be related to "mood, obsessive-compulsive or impulse control disorders." Lejoyeux et al³¹ have linked it to the mood disorders. Some consider CB to be related to the substance

use disorders.^{32,33} Others suggest classifying CB as a disorder of impulse control³⁴ or a mood disorder.³⁵

Faber and O'Guinn²⁶ estimated the prevalence of CB at between 1.8% and 8.1% of the general population, based on results from a mail survey in which the Compulsive Buying Scale (CBS) was administered to 292 individuals selected to approximate the demographic makeup of the general population of Illinois. (The high and low prevalence estimates reflect different score thresholds set for CB.) More recently, Koran et al³⁶ used the CBS to identify compulsive buyers in a random telephone survey of 2513 US adults, and estimated the point prevalence at 5.8% of respondents. Grant et al³⁷ utilized the MIDI to assess CBD and reported a lifetime prevalence of 9.3% among 204 consecutively admitted psychiatric inpatients.

CB has an onset in the late teens/early 20s, which may correlate with emancipation from the nuclear family, as well as with the age at which people can first establish credit.³⁴ Research suggests that 80% to 94% of persons with CBD are women.³⁸ In contrast, Koran et al³⁶ reported that the prevalence of CBD in their random telephone survey was nearly equal for men and women (5.5% and 6.0%, respectively). Their finding suggests that the reported gender difference may be artifactual, in that women more readily acknowledging abnormal shopping behavior than men. Men are more likely to describe their compulsive buying as "collecting."

Data from clinical studies confirm high rates of psychiatric comorbidity, particularly for the mood (21% to 100%), anxiety (41% to 80%), substance use (21% to 46%), and eating disorders (8% to 35%).³⁸ Disorders of impulse control are also relatively common (21% to 40%). The frequency of Axis II disorders in individuals with CB was assessed by Schlosser et al²⁵ using a self-report instrument and a structured interview. Nearly 60% of 46 subjects met criteria for at least one personality disorder through a consensus of both instruments. The most commonly identified personality disorders were the obsessive-compulsive (22%), avoidant (15%), and borderline (15%) types.

A distinctive and stereotyped clinical picture of the compulsive shopper has emerged. Black³⁹ has described four phases including: (i) anticipation; (ii) preparation; (iii) shopping; and (iv) spending. In the first phase, the person with CB becomes preoccupied either with having a specific item, or with the act of shopping. This is followed by a preparation phase in which plans are made. This

Clinical research

phase is followed by the actual shopping experience, which many individuals with CB describe as intensely exciting.²⁵ The act is completed with the purchase, often followed by a sense of let-down or disappointment.³⁶ Perhaps the hallmark of CB is preoccupation with shopping and spending. This typically leads the individual to spend many hours each week engaged in these behaviors.^{24,25} Persons with CB often describe increasing tension or anxiety that is relieved when a purchase is made. CB behaviors occur all year, but can be more problematic during the Christmas season and other holidays, as well as around the birthdays of family members and friends. Compulsive buyers are mainly interested in consumer goods such as clothing, shoes, crafts, jewelry, gifts, makeup, and compact discs (or DVDs)^{24,25} CB has little to do with intellect or educational level, and has been documented in mentally retarded persons.⁴⁰ Similarly, income has relatively little to do with CB, because persons with a low income can be as preoccupied with shopping and spending as wealthier individuals.^{38,40}

Nataraajan and Goff⁴² have identified two independent factors in CB: (i) buying urge or desire, and (ii) degree of control over buying. In their model, compulsive shoppers combine high urge with low control. This view is consistent with clinical reports that compulsive buyers are preoccupied with shopping and spending and will try to resist their urges, often with little success.^{24,38}

Cross-sectional studies suggest the disorder is chronic, though fluctuating in severity and intensity.^{22,25} Aboujaoude et al⁴³ reported that persons who responded to treatment with citalopram were likely to remain in remission during a 1-year follow-up, suggesting that treatment can alter the natural history of the disorder. Lejoyeux et al⁴⁴ report that CB is associated with suicide attempts, although there are no reports of the disorder leading to completed suicide.

There is some evidence that CB runs in families and that within these families mood, anxiety, and substance-use disorders exceed population rates. Black et al⁴⁵ used the family history method to assess 137 first-degree relatives of 31 persons with CB. Relatives were significantly more likely than those in a comparison group to have depression, alcoholism, a drug use disorder, "any psychiatric disorder" and "more than one psychiatric disorder." CB was identified in nearly 10% of the first-degree relatives, but was not assessed in the comparison group.

Neurobiologic theories have centered on disturbed neurotransmission, particularly involving the serotonergic,

dopaminergic, or opioid systems. Selective serotonin reuptake inhibitors (SSRIs) have been used to treat CB,⁴⁶⁻⁵⁰ in part because of hypothetical similarities between CB and OCD, a disorder known to respond to SSRIs. Dopamine has been theorized to play a role in "reward dependence," which has been claimed to foster behavioral addictions, such as CB and PG.¹⁵ Case reports suggesting benefit from the opioid antagonist naltrexone have led to speculation about the role of opioid receptors.⁵¹ There is no direct evidence, however, to support the role of these neurotransmitter systems in the etiology of CB.

Because CB occurs mainly in developed countries, cultural and social factors have been proposed as either causing or promoting the disorder.³⁹ Interestingly, Neuner et al⁵² reported that the frequency of CB in Germany increased following reunification, suggesting that societal factors can contribute to the development of CB. These may include the presence of a market-based economy, the availability of goods, easily obtained credit, and disposable income.¹⁴

There are no standard treatments, and both psychotherapy and medication have been recommended. Several case studies report the psychoanalytic treatment of CB.⁵³⁻⁵⁵ More recently, cognitive-behavioral treatment (CBT) models have been developed for CB, many of them employing group therapy.^{56,57} Mitchell et al⁵⁷ found that group CBT produced significant improvement compared with a wait-list in a 12-week pilot study. Improvement attributed to CBT was maintained during a 6-month follow-up. Benson⁵⁸ has developed a comprehensive self-help program that can be used by both individuals and groups.

Treatment studies employing psychotropic medications have produced mixed results. Early reports suggested the benefit of antidepressants in treating CB.^{22,23} Black et al⁴⁶ reported the results of an open-label trial in which subjects given fluvoxamine showed benefit. Two subsequent randomized controlled trials (RCTs) found fluvoxamine treatment to be no better than placebo.^{47,48} Koran et al⁵¹ later reported that subjects with CB improved with open-label citalopram. In a subsequent study, subjects received open-label citalopram; those who were considered responders were randomized to citalopram or placebo. Compulsive shopping symptoms returned in 5/8 subjects (62.5%) assigned to placebo compared with 0/7 who continued taking citalopram. In an identically designed discontinuation trial, escitalopram did not separate from placebo.⁵² Because the medication study findings are mixed, no empirically well-sup-

ported treatment recommendations can be made. Open-label trials have generally produced positive results, but RCTs have not. Interpretation of these study results is complicated by placebo response rates as high as 64%.⁴⁷

Pathological gambling

PG is increasingly being recognized as a major public health problem.⁵⁹ PG is estimated to cost society approximately \$5 billion per year and an additional \$40 billion in lifetime costs for reduced productivity, social services, and creditor losses. The disorder substantially impairs quality of life in addition to its association with comorbid psychiatric disorders, psychosocial impairment, and suicide.^{59,61} Family-related problems include financial distress, child and spousal abuse, and divorce and separation.⁶¹

While problematic gambling behavior has been recognized for centuries, it was often ignored by the psychiatric community. Bleuler,¹⁷ citing Kraepelin,¹⁶ considered PG, or "gambling mania," a *special impulse* disorder. Criteria for PG were first enumerated in 1980 in *DSM-III*.⁶² The criteria were subsequently modified, and in *DSM-IV-TR*,¹⁰ are patterned after those used for substance dependencies and emphasize the features of tolerance and withdrawal. PG is defined as "persistent and recurrent maladaptive gambling behavior (criterion A) that disrupts personal, family, or vocational pursuits..." Ten specific maladaptive behaviors are listed, and ≥5 are required for the diagnosis. The criteria focus on loss of control of gambling behavior; progressive deterioration of the disorder; and continuation despite negative consequences. The diagnosis can only be made when mania is ruled out (Criterion B). In an attempt to reconcile nomenclature and measurement methods, Shaffer and Hall⁶³ developed a generic multilevel classification scheme that is now widely accepted by gambling researchers.

PG is presently classified as a disorder of impulse control in *DSM-IV-TR*.¹⁰ On the one hand, some investigators have suggested that PG is related to OCD,^{1,64} yet others argue against such a relationship.⁶⁵ On the other hand, PG is widely considered an addictive disorder.^{66,67} It has recently been proposed as a candidate for inclusion in a new category for "behavioral addictions."¹⁵ Recent estimates of lifetime prevalence for PG range from 1.2% to 3.4% in the general population.^{68,69} Prevalence rates have risen in areas where gambling availability has increased.^{70,71} A national survey showed

that the availability of a casino within 50 miles is associated with a nearly twofold increase in PG prevalence.⁵⁹ Gambling behavior typically begins in adolescence, with PG developing by the late 20s or early 30s,⁷² though it can begin at any age through senescence. Rates of PG are higher in men, but the gender gap may be narrowing. PG has a later onset in women yet progresses more rapidly ("telescoping") than in men,⁷³ at a rate similar to that observed in alcohol disorders. Populations at risk include adults with mental health or substance-use disorders, persons who have been incarcerated, African-Americans, and persons with low socioeconomic status.^{74,75}

Research has not validated PG subtypes, but perhaps the most widely discussed distinction is between "escape-seekers" and "sensation-seekers."⁷⁶ Escape-seekers are often older persons who gamble out of boredom, from depression, or to fill time, and choose passive forms of gambling such as slot machines. Sensation-seekers tend to be younger, and prefer the excitement of card games or table games that involve active input.⁷⁶ Blaszczynski and Nower⁷⁷ have proposed a "pathways" model that integrates biological, developmental, cognitive, and other determinants of disordered gambling. They have identified three subgroups: a) behaviorally-conditioned gamblers; b) emotionally vulnerable gamblers; and c) antisocial, impulsive gamblers. Behaviorally conditioned gamblers have no specific predisposing psychopathology, but make bad judgments regarding gambling. Emotionally vulnerable gamblers suffer premorbid depression or anxiety, and have a history of poor coping. Finally, antisocial, impulsive gamblers are highly disturbed and have features of antisocial personality disorder and impulsivity that suggest neurobiological dysfunction.

Psychiatric comorbidity is the rule, not the exception, in persons with PG. Both community and clinic-based studies suggest that substance use disorders, mood disorders, and personality disorders are highly prevalent in persons with PG.⁷⁸ In clinical samples, from 25% to 63% of pathological gamblers meet lifetime criteria for a substance use disorder.⁷⁹ Correspondingly, from 9% to 16% of substance abusers are probable pathological gamblers.⁷⁹ PG is also associated with increased prevalence of mood disorders, and overall 13% to 78% of persons with pathological gambling are estimated to experience a mood disorder.⁷⁹ On the other hand, patients with mood disorders have not been found to have elevated rates of PG. Rates of other impulse-control disorders (ICDs) appear higher in persons with pathological gambling than in the

Clinical research

general population. Investigators have reported rates ranging from 18% to 43% for one or more ICD.⁷⁹ CB appears to be the most frequent comorbid ICD in persons with PG, perhaps because both disorders share characteristics of focused attention, monetary gratification, and monetary exchange. Subjects with one ICD appear more likely to have another, suggesting considerable overlap among them.

Personality disorders are relatively common among individuals with PG, particularly those in “cluster B.” Antisocial personality disorder has been singled out as having a close relationship with PG, perhaps because crime and gambling frequently co-occur, with rates ranging from 15% to 40%.^{79,80} At least one study of persons with antisocial personality disorder showed high rates of PG.⁸¹

PG is widely thought to be chronic and progressive.^{82,83} This view is embedded in *DSM-IV-TR*¹⁰ which holds that the essential feature of PG is “persistent and recurrent maladaptive gambling behavior ...that disrupts personal, family, or vocational pursuits” (p 671). These views were influenced by the pioneering observations of Custer⁸⁴ who described PG as a progressive, multistage illness that begins with a *winning phase*, followed in turn by a *losing phase*, and a *desperation phase*. The final phase, *giving up*, represented feelings of hopelessness.⁸⁵ Some contend that many pathological gamblers experience a “big win” early in their gambling careers that leads directly to their becoming addicted. Custer’s four phases of PG have gained wide acceptance despite the absence of empirical data.

Recent work is leading to a reconsideration of these views. LaPlante et al⁸⁶ reviewed five studies⁸⁷⁻⁹¹ that met their criteria of reporting longitudinal data pertaining to gambling that did not involve a treatment sample. LaPlante et al report that, from the four studies that included level 3 gamblers (ie, persons with PG), most gamblers improved, and moved to a lower level, and that rates of classification improvement were “at least significantly greater than 29%.” Results were similar for level 2 (ie, “at-risk”) gamblers. Those who were level 0 to 1 gamblers at baseline were unlikely to progress to a higher (ie, more severe) level of gambling behavior, and with one exception,⁹¹ the studies suggested that few level 2 gamblers improved by moving to level 1. La Plante et al⁸⁶ conclude that these studies challenge the notion that PG is intractable, and suggest that many gamblers spontaneously improve, as do many substance addicted persons. The findings suggest that those who do not gamble

or gamble without problems tend to remain problem-free; those with disordered gambling move from one level to another, though the general direction is toward improved classification.

Family history data suggests that PG, mood disorders, and substance-use disorders are more prevalent among the relatives of persons with PG than in the general population.^{92,93} Twin studies also suggest that gambling has a heritable component.⁹⁴ Functional neuroimaging studies suggest that among persons with PG, gambling cues elicit gambling urges and a temporally dynamic pattern of brain activity changes in frontal, paralimbic, and limbic brain structures, suggesting to some extent that gambling may represent dysfunctional frontolimbic activity.⁹⁵ There is little consensus about the appropriate treatment of PG. Few persons with PG seek treatment,⁹⁶ and until recently the treatment mainstay appeared to be participation in Gamblers Anonymous (GA), a 12-step program patterned after Alcoholics Anonymous. Attendance at GA is free and chapters are available throughout the US, but follow-through is poor and success rates disappointing.⁹⁷ Inpatient treatment and rehabilitation programs similar to those for substance-use disorders have been developed, and are helpful to some^{98,99} Still, these programs are unavailable to most persons with PG because of geography or lack of access (ie, insurance/financial resources). More recently, CBT and motivational interviewing have become established treatment methods.¹⁰⁰ Self-exclusion programs have also gained acceptance and appear to benefit selected patients.¹⁰¹ While rules vary, they generally involve voluntary self-exclusion from casinos for a period of time at the risk of being arrested for trespassing. Medication treatment studies have gained momentum, but their results are inconsistent. Briefly, the opioid antagonists naltrexone and nalmefene were superior to placebo in randomized controlled trials (RCTs)^{102,103} but controlled trials of paroxetine and bupropion were negative.^{104,105} Open-label studies of nefazodone, citalopram, carbamazepine, and escitalopram have been encouraging, but need to be followed up with adequately powered and controlled studies.¹⁰⁶⁻¹⁰⁹

Putative relationship between CB/PG and OCD

The relationship between CB/PG and OCD remains uncertain. The inclusion of CB and PG within an OC spectrum, while intriguing, rests on hypothesis and not

empirical data. How these disorders should be classified has been debated for nearly 100 years. Opinion has mainly favored their inclusion among disorders of impulse control. For historical reasons, and because of the lack of empirical data, we believe that the two disorders should remain with the ICDs until convincing evidence is presented to favor their inclusion either with the addictive disorders or an OC spectrum.

The most obvious connection between CB and PG and OCD is phenomenologic. Each disorder involves repetitive behavior that generally occurs in response to overwhelming thoughts and urges; engaging in the behavior—at least temporarily—will satisfy the urge, and/or reduce tension and anxiety that preceded the behavior. Nonetheless, a fundamental distinction between CB/PG and OCD is that the behaviors (shopping, gambling) are considered *ego-syntonic*; that is, they are viewed as pleasurable and desirable, while behaviors associated with OCD never are, and nearly all patients want to be rid of them. Not so with shopping and gambling: the person with CB or PG finds the behaviors highly pleasurable, and only wants to stop the behaviors when their deleterious secondary consequences become overwhelming. Proponents of the OC spectrum point to the overlap between these disorders and OCD. Comorbidity studies have found that in clinical samples from 3% to 35% of individuals with CB have comorbid OCD.^{22,46} In fact, the presence of CB may characterize a specific subset of OCD patients,^{110,111} particularly those who hoard. Hoarding is a special symptom that involves the acquisition of and failure to discard, possessions that are of limited use or value.¹¹² Yet, unlike the items retained by the typical hoarder, the items purchased by the person with CB are not inherently valueless or useless.

CB frequently appears to be comorbid with the ICDs. Black and Moyer⁸⁰ and Grant and Kim⁷² each reported elevated rates of CB among samples of pathological gamblers (23% and 8%, respectively). Likewise, other impulse control disorders are common among compulsive shoppers.³⁹ Comorbidity studies of PG are more mixed, although they generally report higher rates of OCD than in the general population. The reverse does not seem to be true. Axis II comparisons show that the predominant disorders associated with OCD are the “cluster C” disorders. While there are no axis II disorders specifically associated with PG or CB, “cluster B” disorders appear overrepresented, particularly antisocial personality disorder.

Direct investigations into OC characteristics of persons with PG found that those with PG scored higher than those without on scales measuring OC traits.⁶⁴ CB and PG also share high trait impulsivity.^{19,113}

Other evidence could come from family studies of CB, PG, or OCD. There are few family studies regarding these disorders, and none have supported a familial relationship among these disorders. In the only controlled family history study of CB, Black et al⁴⁵ did not find a relationship with OCD. In two family studies, one using the family history method, the other using the family interview method, the investigators were unable to establish a connection between PG and OCD.^{114,115} Looking at this connection through OCD family studies has also failed to find a connection. Neither Black et al¹¹⁴ nor Bienvenu et al¹¹⁵ were able to establish a familial relationship between OCD and PG.

Demographic similarities are often used to suggest that disorders might be linked, for example the fact that both alcohol disorders and antisocial personality disorder are predominantly found in men. Yet, there is no similarity in gender distribution among these disorders. With PG there is a clear male preponderance; with CB a female preponderance; with OCD, the gender distribution is evenly split.

If these disorders were related, their natural history and course might be similar as well. CB and OCD appear to have an onset in the late teens or early 20s. PG appears to have a slightly later onset, with women developing the disorder much later than men, but having a briefer course from onset of gambling to development of a disorder. This is what is seen with alcohol disorders, but not OCD. With CB, PG, and OCD are all considered mostly chronic, but the similarity stops there. For CB and PG, while there are no careful, longitudinal studies, the data suggest that the disorders may be episodic, that is, may remit for varying lengths of time depending on a host of external factors such as fear of consequences, eg, bankruptcy or divorce, or lack of income; OCD rarely remits. In terms of suicide risk, PG has been reported to carry a risk for suicide attempts and completed suicides; with CB, there are anecdotal reports of suicide attempts, but not completed suicides; with OCD, the data is somewhat mixed, but overall, the risk of completed suicide is considered low.

Here, too, when one considers treatment response, OCD is well known to respond well to serotonin reuptake inhibitor antidepressants, and to cognitive behavioral therapy. CB and PG have no clear response to medica-

Clinical research

tion, and the most robust treatment data suggests that PG may respond to opioid antagonists. Both CB and PG are reported to respond to CBT, but the completeness and quality of the response is unlike that seen with OCD.

The presence of similar biological markers is another way to assess the connection between these disorders. This task is hampered by the fact that none of these disorders has reliable markers. Nonetheless, a functional magnetic resonance imaging (fMRI) study of PG suggests that the disorder shows an abnormal pattern of activation in specific subcortical-frontal regions following cue exposure. Potenza et al⁸⁶ interpret these findings as evidence for the similarity of brain pathways in PG and drug addiction, while the opposite direction of higher brain activation is found in OCD. Similarly, Goodriaan et al¹¹⁶ review the research on neurochemical and molecular genetic data involving PG. They conclude that there is evidence of disturbed neurotransmission involving dopamine (DA), serotonin, and norepinephrine; and “... are in accordance with the findings of abnormal brain activation in reward pathways, where DA is an important transmitter” (p 134). Dopamine is noted to play an important role in craving and withdrawal in the substance use disorders. While the neurotransmission involved in OCD has not been fully elucidated, the central serotonin system has been the most actively studied. This is perhaps due to the robust effect of SSRIs in the treatment of OCD.

On the whole, neuropsychological studies of PG indicate that pathological gamblers have impaired performance in several aspects of executive function including attention, delay discounting, and decision-making.¹¹⁵⁻¹¹⁷ With OCD, neuropsychological research is less consistent; there is evidence of impaired response-inhibition and in attentional set-shifting, but little evidence of impaired reversal learning and decision-making.¹¹⁸ To our knowledge, there are no neuropsychological studies of persons with CB.

Alternate classification schemes

If CB and PG are not part of an OC spectrum, where should they be classified? Because there is almost no evidence suggesting a relationship with the mood disorders, that possibility can probably be eliminated outright. Of the remaining schemes, the most likely candidates are to include PG and CB with the ICDs, or to move them to a category involving the substance-use disorders. Keeping PG and CB with the ICDs is the easiest option: PG is already classified as an ICD, and while CB is not

currently included in *DSM-IV-TR*, it has historically been considered an impulsive disorder. Both PG and CB share similar clinical features involving the presence of irresistible, ego-syntonic urges that prompt a behavioral response. The response (ie, gambling, shopping) satisfies the urge and/or temporarily reduces tension or anxiety, but is often followed by a sense of guilt or shame, and ultimately leads to adverse, secondary consequences. The behaviors are chronic or intermittent, and may spontaneously remit, sometimes in response to external circumstances. Age of onset and gender distribution differ, as discussed earlier. Possibly, CB may be considered the female equivalent of PG, because they tend to have a reverse gender distribution: men predominate among those with PG; women predominate among those with CB. Both appear to respond to CBT, yet neither has a clear response to medication; SSRIs do not produce consistent improvement. Comorbidity studies show overlap among the disorders, as a disproportionate number of pathological gamblers have CB and vice versa.

On the other hand, data suggest many commonalities with the substance use disorders. PG and CB are both associated with cravings that are not unlike those reported by substance abusers; PG is noted to produce “withdrawal” symptoms when the gambler is abstinent,¹¹⁹ though this has not been studied in CB. Research shows that persons with PG or CB often have comorbid substance use disorders. Conversely, substance abusers have high rates of PG; there are no comparable data for CB. Family studies show that relatives of probands with PG or CB have high rates of psychiatric illness, particularly alcohol and drug use disorders. Further, Slutske et al⁹⁴ have reported that, based on twin data, PG appears to be related to the substance-use disorders and antisocial personality disorder. Finally, as noted earlier, the neuroimaging studies, and both neurotransmitters and molecular genetic research on PG suggest a relationship with the substance-use disorders.¹¹⁶ These data support the inclusion of PG and perhaps CB in a category for “behavioral addictions,” possibly comprising a subset of the substance-use disorders, but they do not support a relationship with OCD.

Conclusions

The review suggests that CB and PG are probably not candidates for inclusion in an OC spectrum. The review was not meant to judge the merit of the OC spectrum concept.

In fact, we have suggested that there appears to be sufficient evidence to support the existence of a limited OC spectrum that might include body dysmorphic disorder, Tourette's disorder, trichotillomania, subclinical OCD, and perhaps the grooming disorders.^{8,120} While there are superficial phenomenologic similarities between CB/PG and OCD, other evidence suggests they are not associated: gender distribution, age at onset, and course; comorbidity

studies; neuroimaging, neurotransmitter, and neuropsychological studies; and treatment response. We believe that PG and CB are likely related, despite their much different gender distribution. Further, we believe that in the absence of new and convincing evidence, PG ought to remain within the ICD category. Lastly, we believe that CB is an identifiable and distinct disorder that ought to be included in *DSM-5*, and should be included with the ICDs. □

REFERENCES

1. Hollander E. *Obsessive Compulsive Related Disorders*. Washington DC: American Psychiatric Press; 1993.
2. Hollander E. Obsessive-compulsive spectrum disorders: an overview. *Psychiatr Ann*. 1993;23:355-358.
3. Hollander E, Wong CM. Introduction: obsessive-compulsive spectrum disorders. *J Clin Psychiatry*. 1995;56(suppl 4):3-6.
4. Koran LM. *Obsessive-Compulsive and Related Disorders in Adults – a Comprehensive Clinical Guide*. New York NY; Cambridge, UK: 1999.
5. Rasmussen SA. Obsessive-compulsive spectrum disorders. *J Clin Psychiatry*. 1994;55:89-91.
6. Castle DJ, Phillips KA. Obsessive-compulsive spectrum of disorders: a defensible construct? *Aust NZ J Psychiatry*. 2006;40:114-120.
7. Tavares H, Gentil V. Pathological gambling and obsessive compulsive disorder: towards a spectrum of disorders of volition. *Rev Brasil Psiquiatria*. 2007;29:107-117.
8. Black DW. The obsessive-compulsive spectrum: fact or fancy? In: Maj M, Sartorius N, Okasha A, Zohar J, eds. *Obsessive-Compulsive Disorder*. New York, NY: Wiley; 2000:233-235.
9. Phillips KA. The obsessive-compulsive spectrum: promises and pitfalls. In: Maj M, Sartorius N, Okasha A, Zohar J, eds. *Obsessive-Compulsive Disorder*. New York, NY: Wiley; 2000:225-227.
10. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th ed, Text Revision. Washington, DC: American Psychiatric Association; 2000.
11. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *Am J Psychiatry*. 1970;126:983-987.
12. World Health Organization. *International Classification of Diseases*. 9th Revision. Geneva, Switzerland: World Health Organization; 1977.
13. Zohar J. The Cape Town Consensus Group Consensus Statement for Obsessive-Compulsive Spectrum to Obsessive Compulsive Disorder: The Cape Town Consensus Statement. *CNS Spectr* 2007;12:2(suppl 3): 5-13.
14. Lee S, Mysyk A. The medicalization of compulsive buying. *Soc Sci Med*. 2004;58:1709-1718.
15. Holden C. Behavioral addictions: so they exist? *Science*. 2001;294: 980-982.
16. Kraepelin E. *Psychiatrie*. 8th ed. Leipzig, Germany: Verlag Von Johann Ambrosius Barth; 1915:408-409.
17. Bleuler E. *Textbook of Psychiatry*. AA Brill, Trans. New York, NY: Macmillan; 1930.
18. Esquirol JED. *Des maladies mentales*. Paris, France: Baillière; 1838.
19. O'Guinn TC, Faber RJ. Compulsive buying: a phenomenological exploration. *J Consumer Res*. 1989;16:147-157.
20. Elliott R. Addictive consumption: function and fragmentation in post-modernity. *J Consumer Policy*. 1994;17:159-179.
21. Magee A. Compulsive buying tendency as a predictor of attitudes and perceptions. *Adv Consum Res*. 1994;21:590-594.
22. McElroy S, Keck PE, Pope HG, et al. Compulsive buying: a report of 20 cases. *J Clin Psychiatry*. 1994;55:242-248.
23. McElroy S, Satlin A, Pope HG, et al. Treatment of compulsive shopping with antidepressants: a report of three cases. *Ann Clin Psychiatry*. 1991;3:199-204.
24. Christenson GA, Faber RJ, de Zwaan M, et al. Compulsive buying: descriptive characteristics and psychiatric comorbidity. *J Clin Psychiatry*. 1994;55:5-11.
25. Schlosser S, Black DW, Repertinger S, Freet D. Compulsive buying: demography, phenomenology, and comorbidity in 46 subjects. *Gen Hosp Psychiatry*. 1994;16:205-212.
26. Faber RJ, and O'Guinn TC. A clinical screener for compulsive buying. *J Consumer Res*. 1992;459-469.
27. Edwards EA. Development of a new scale to measure compulsive buying behavior. *Fin Counsel Plan*. 1993;4:67-84.
28. Dittmar H. Understanding and diagnosing compulsive buying. In: Coombs R, ed. *Addictive Disorders. A Practical Handbook*. New York, NY: Wiley; 2004:411-450.
29. Budden MC, Griffin TF. Explorations and implications of aberrant consumer behavior. *Psychol Marketing*. 1996;13:739-740.
30. Hollander E, Allen A. Is compulsive buying a real disorder and is it really compulsive? *Am J Psychiatry*. 2006;163:1670-1672.
31. Lejoyeux M, Andes J, Tassian V, Solomon J. Phenomenology and psychopathology of uncontrolled buying. *Am J Psychiatry*. 1996;152:1524-1529.
32. Glatt MM, Cook CC. Pathological spending as a form of psychological dependence. *Br J Addict*. 1987;82:1252-1258.
33. Goldman R. Compulsive buying as an addiction. In: Benson A, ed. *I Shop, Therefore I Am: Compulsive Buying and the Search For Self*. New York, NY: Jason Aronson; 2000:245-267.
34. Black DW. Compulsive buying disorder: definition, assessment, epidemiology and clinical management. *CNS Drugs*. 2001;15:17-27.
35. McElroy SE, Pope HG, Keck PE, et al. Are impulse control disorders related to bipolar disorder? *Compr Psychiatry*. 1996;37:229-240.
36. Koran LM, Faber RJ, Aboujaoude E, et al. Estimated prevalence of compulsive buying in the United States. *Am J Psychiatry*. 2006;163:1806-1812.
37. Grant JE, Levine L, Kim SW, Potenza MN. Impulse control disorders in adult psychiatric inpatients. *Am J Psychiatry*. 2005;162:2184-2188.
38. Black DW. Epidemiology and phenomenology of compulsive buying disorder. In: Grant J, Potenza M, eds. *Oxford Handbook of Impulse Control Disorders*. In press.
39. Black DW. Compulsive buying disorder: a review of the evidence. *CNS Spectrums*. 2007;12:124-132.
40. Otter M, Black DW. Compulsive buying behavior in two mentally challenged persons. *Prim Care Companion J Clin Psychiatry*. 2007;9:469-470.
41. Dittmar H. When a better self is only a button click away: associations between materialistic values, emotional and identity-related buying motives, and compulsive buying tendency online. *J Soc Clin Psychol*. 2007;26:334-361.
42. Natarajan R, Goff BG. Compulsive buying: toward a reconceptualization. *J Soc Behav Person*. 1991;6:307-328.
43. Aboujaoude E, Gamel N, Koran LM. A 1-year naturalistic follow-up of patients with compulsive shopping disorder. *J Clin Psychiatry*. 2003;64:946-950.
44. Lejoyeux M, Tassian V, Solomon J, Ades J. Study of compulsive buying in depressed persons. *J Clin Psychiatry*. 1997;58:169-173.
45. Black DW, Repertinger S, Gaffney GR, Gabel J. Family history and psychiatric comorbidity in persons with compulsive buying: preliminary findings. *Am J Psychiatry*. 1998;155:960-963.

Clinical research

El juego patológico y el comprar compulsivo: ¿corresponde incluirlos dentro del espectro obsesivo-compulsivo?

Se ha propuesto que el comprar compulsivo (CC) y el juego patológico (JP) se integren en el espectro de los trastornos relacionados con el trastorno obsesivo compulsivo (TOC). La hipótesis del espectro se originó a comienzos de la década de 1990 y ha conseguido bastante apoyo, a pesar de la falta de evidencias empíricas. El interés en esta hipótesis ha llegado a un punto crítico ya que algunos investigadores han recomendado la creación de una nueva categoría que incluya estos trastornos en el DSM-V, que está actualmente en desarrollo. En este artículo los autores describen el origen del espectro obsesivo-compulsivo (OC) y sus fundamentos teóricos, revisan el CC y el JP, y discuten los datos a favor y en contra de un espectro OC. Ambos trastornos son descritos en términos de su historia, definición, clasificación, fenomenología, historia familiar, fisiopatología y manejo clínico. Los autores concluyen que: 1) el CC y el JP probablemente no se relacionan con el TOC y no es suficiente la evidencia para incluirlos en el espectro OC dentro del DSM-V, 2) el JP debiera incluirse dentro de los trastornos del control impulsivo (TCI) y 3) se debe crear un nuevo diagnóstico del CC y clasificarlo como un TCI.

Jeu pathologique et achat compulsif : font-ils partie du spectre des troubles obsessionnels-compulsifs ?

Certains auteurs ont proposé d'intégrer l'achat compulsif (AC) et le jeu pathologique (JP) dans le spectre des troubles obsessionnels-compulsifs (TOC), concept émergeant au début des années 90, et qui a reçu un soutien important en dépit d'un manque de preuves empiriques. L'intérêt pour cette hypothèse est devenu très important en raison de la recommandation de certains experts de créer une nouvelle catégorie incluant ces troubles dans le DSM-5 actuellement en rédaction. Dans cet article, les auteurs décrivent l'origine des troubles obsessionnels-compulsifs (TOC) et de leurs bases théoriques, analysent le JP et l'AC et examinent les arguments pour et contre leur appartenance au spectre des TOC. Les deux pathologies sont décrites en termes d'historique, de définition, de classification, de phénoménologie, d'antécédents familiaux, de physiopathologie et de prise en charge clinique. Les auteurs concluent que : (i) le JP et l'AC ne sont probablement pas liés aux TOC et que les preuves sont insuffisantes pour les placer dans le cadre OC du DSM-V ; (ii) le JP devrait rester au sein des troubles du contrôle de l'impulsion (TCI) ; et (iii) une nouvelle définition de l'AC devrait être créée pour le classer également dans les TCI.

46. Black DW, Monahan P, Gabel J. Fluvoxamine in the treatment of compulsive buying. *J Clin Psychiatry*. 1997;58:159-163.
47. Black DW, Gabel J, Hansen J, et al. A double-blind comparison of fluvoxamine versus placebo in the treatment of compulsive buying disorder. *Ann Clin Psychiatry*. 2000;12:205-211.
48. Ninan PT, McElroy SL, Kane CP, et al. Placebo-controlled study of fluvoxamine in the treatment of patients with compulsive buying. *J Clin Psychopharmacol*. 2000;20:362-366.
49. Koran LM, Chuang HW, Bullock KD, Smith SC. Citalopram for compulsive shopping disorder: an open-label study followed by a double-blind discontinuation. *J Clin Psychiatry* 2003;64:793-798.
50. Koran LM, Aboujaoude EN, Solvason B, Gamel N, Smith EH. Escitalopram for compulsive buying disorder: a double-blind discontinuation study. *J Clin Psychopharmacol*. 2007;27:225-227. Letter.
51. Grant JE. Three cases of compulsive buying treated with naltrexone. *Int J Psychiatry Clin Prac*. 2003;7:223-225.
52. Neuner M, Raab G, Reisch L. Compulsive buying in maturing consumer societies: an empirical re-inquiry. *J Econ Psychol*. 2005;26:509-522.
53. Krueger DW. On compulsive shopping and spending: a psychodynamic inquiry. *Am J Psychother*. 1988;42:574-584.
54. Lawrence L. The psychodynamics of the compulsive female shopper. *Am J Psychoanal*. 1990;50:67-70.
55. Winestine MC. Compulsive shopping as a derivative of childhood seduction. *Psychoanal Q*. 1985;54:70-72.
56. Villarino R, Otero-Lopez JL, Casto R. *Adicion a la compra: Analysis, evaluation y tratamiento [Buying addiction: Analysis, evaluation, and treatment]*. Madrid, Spain: Ediciones Pirámide; 2001.
57. Mitchell JE, Burgard M, Faber R, Crosby RD. Cognitive behavioral therapy for compulsive buying disorder. *Behav Res Ther*. 2006;44:1859-1865.
58. Benson A. *Stopping Overshopping – a Comprehensive Program to Help Eliminate Overshopping*. New York, NY: April Benson; 2006.
59. National Opinion Research Center at the University of Chicago (NORC): Gambling Impact and Behavior Study, Report to the National Gambling Impact Study Commission, April 1, 1999.
60. Petry NM, Kiluk BD. Suicidal ideation and suicide attempts in treatment-seeking pathological gamblers. *J Nerv Ment Dis*. 2002;190:462-469.
61. Shaw M, Forbush K, Schlinger J, et al. The effect of pathological gambling on families, marriages, and children. *CNS Spectr*. 2007;12:615-622.
62. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 3rd ed. Washington, DC: American Psychiatric Association; 1980.
63. Shaffer HJ, Hall MN. Estimating the prevalence of adolescent gambling disorders: a quantitative synthesis and guide toward standard gambling nomenclature. *J Gambl Stud*. 1996;12:193-214.
64. Blaszczynski A. Pathological gambling and obsessive-compulsive spectrum disorders. *Psychol Rep*. 1999;84:107-113.
65. Durdle H, Gorey KM, Stewart SH. A meta-analysis examining the relations among pathological gambling, obsessive-compulsive disorder, and obsessive-compulsive traits. *Psychol Rep*. 2008;103:485-498.

- 66.** Shaffer HJ, LaPlante DA, LaBrie RA, et al. Toward a syndrome model of addiction: multiple expressions, common etiology. *Harv Rev Psychiatry*. 2004;12:367-374.
- 67.** Wray I, Dickerson MG. Cessation of high frequency gambling and withdrawal symptoms. *Br J Addiction*. 1981;76:401-405.
- 68.** Shaffer HJ, Hall MN. Updating and refining prevalence estimates of disordered gambling behavior in the United States and Canada. *Can J Pub Health*. 2001;92:168-172.
- 69.** Cunningham-Williams R, Cottler LB. The epidemiology of pathological gambling. *Sem Clin Neuropsychiatry*. 2001;6:155-166.
- 70.** Volberg RA. Prevalence studies of problem gambling in the United States. *J Gambling Stud*. 1996;12:111-128.
- 71.** Jacques C, Ladouceur R, Gerland F. Impact of availability on gambling: a longitudinal study. *Can J Psychiatry*. 2000;45:810-815.
- 72.** Grant J, Kim SW. Demographic and clinical features of 131 adult pathological gamblers. *J Clin Psychiatry*. 2001;62:957-962.
- 73.** Tavares H, Zilberman ML, Beites FJ, et al. Gender differences in gambling progression. *J Gambl Stud*. 2001;17:151-159.
- 74.** Potenza MN, Kosten TR, Rounsaville BJ. Pathological gambling. *JAMA*. 2001;286:141-144.
- 75.** Templer DI, Kaiser G, Siscoe K. Correlates of pathological gambling propensity in prison inmates. *Compr Psychiatry*. 1993;34:347-351.
- 76.** Blaszczynski A, McConaghy N. Anxiety and/or depression in the pathogenesis of addictive gambling. *Int J Addictions*. 1989;24:337-350.
- 77.** Blaszczynski A, Nower L. A pathways model of problem and pathological gambling. *Addiction*. 2002;97:487-499.
- 78.** Crockford ND, el-Guebaly N. Psychiatric comorbidity in pathological gambling: a critical review. *Am J Psychiatry*. 1998;143:43-50.
- 79.** Black DW, Shaw M. Psychiatric comorbidity and pathological gambling. *Psychiatric Times*. 2008;25:14-18.
- 80.** Black DW, Moyer T. Clinical features and psychiatric comorbidity in 30 subjects reporting pathological gambling behavior. *Psychiatr Serv*. 1998;49:1434-1439.
- 81.** Goldstein RB, Powers SI, McCusker J, et al. Lack of remorse in antisocial personality disorder among drug abusers in residential treatment. *J Pers Disord*. 1996;10:321-334.
- 82.** Cartwright C, DeCaria C, Hollander E. Pathological gambling: a clinical review. *J Prac Psychiatr Behav Health*. 1998;5:277-286.
- 83.** DeCaria C, Hollander E, Grossman R, et al. Diagnosis, neurobiology, and treatment of pathological gambling. *J Clin Psychiatry*. 1996;57(suppl 8):80-84.
- 84.** Custer R. *When Luck Runs Out*. New York, NY: Facts on File; 1985:232.
- 85.** Rosenthal R. Pathological gambling. *Psychiatr Ann*. 1992;22:72-78.
- 86.** LaPlante DA, Nelson SE, LaBrie RA, Shaffer HJ. Stability and progression of disordered gambling: lessons from longitudinal studies. *Can J Psychiatry*. 2008;53:52-60.
- 87.** Abbott MW, Williams MM, Volberg RA. A prospective study of problem and regular non-problem gamblers living in the community. *Subst Use Misuse*. 2004;39:855-884.
- 88.** DeFuentes-Merillas L, Koeter MW, Schippers GM, van den Brink W. Temporal stability of pathological scratchcard gambling among adult scratchcard buyers two years later. *Addiction*. 2004;99:117-127.
- 89.** Shaffer HJ, Hall MN. The natural history of gambling and drinking problems among casino workers. *J Soc Psychol*. 2002;142:405-424.
- 90.** Slutsko W, Jackson KM, Sher KJ. The natural history of problem gambling from age 18 to 29. *J Abnorm Psychol*. 2003;112:263-274.
- 91.** Winters KC, Stinchfield RD, Botzett A, Anderson N. A prospective study of youth gambling behaviors. *Psychol Addict Behav*. 2002;16:3-9.
- 92.** Black DW, Moyer T, Schlosser S. Quality of life and family history in pathological gambling. *J Nerv Ment Dis*. 2003;191:124-126.
- 93.** Black DW, Monahan PO, Temkit M, Shaw M. A family study of pathological gambling. *Psychiatr Res*. 2006;141:295-303.
- 94.** Slutsko W, Eisen S, True WR, et al. Common genetic vulnerability for pathological gambling and alcohol dependence in men. *Arch Gen Psychiatry*. 2000;57:666-673.
- 95.** Potenza MN, Steinberg MA, Skudlarski P, et al. Gambling urges and pathological gambling: a functional magnetic resonance imaging study. *Arch Gen Psychiatry*. 2003;60:828-836.
- 96.** Cunningham JA. Little use of treatment among problem gamblers. *Psychiatr Serv*. 2005;56:1024-1025.
- 97.** Brown RIF. The Effectiveness of Gamblers Anonymous. In Edington WR (ed), *The Gambling Studies: Proceedings of the Sixth National Conference on Gambling and Risk Taking*. Reno, NV: Bureau of Business and Economic Research, University of Nevada, Reno; 1985.
- 98.** Russo AM, Taber JL, McCormick RA, Ramirez LF. An outcome study of an inpatient program for pathological gamblers. *Hosp Comm Psychiatry*. 1984;35:823-827.
- 99.** Taber JL, McCormick RA, Russo AM, et al. A follow-up of pathological gamblers after treatment. *Am J Psychiatry*. 1987;144:757-761.
- 100.** Petry NM. *Pathological Gambling: Etiology, Comorbidity, and Treatment*. Washington DC: American Psychological Association, 2005.
- 101.** Ladouceur R, Sylvain C, Gosselin P. Self-exclusion program: a longitudinal evaluation study. *J Gambl Stud*. 2007;23:85-94.
- 102.** Kim SW, Grant JE, Adson DE, Shin YC. Double-blind naltrexone and placebo comparison study in the treatment of pathological gambling. *Biol Psychiatry*. 2001;49:914-921.
- 103.** Grant JE, Potenza MN, Hollander E, et al. Multicenter investigation of the opioid antagonist nalmefene in the treatment of pathological gambling. *Am J Psychiatry*. 2006;163:303-312.
- 104.** Black DW, Arndt S, Coryell WH, et al. Bupropion in the treatment of pathological gambling: a randomized, placebo-controlled, flexible-dose study. *J Clin Psychopharmacol*. 2007;27:143-150.
- 105.** Grant JE, Potenza MN, Blanco C, et al. Paroxetine treatment of pathological gambling: a multi-center randomized controlled trial. *Int Clin Psychopharmacol*. 2003;18:243-249.
- 106.** Pallanti S, Rossi NB, Sood E, Hollander E. Nefazodone treatment of pathological gambling: a prospective open-label controlled trial. *J Clin Psychiatry*. 2002;63:1034-1039.
- 107.** Zimmerman M, Breen RB, Posternak MA. An open-label study of citalopram in the treatment of pathological gambling. *J Clin Psychiatry*. 2002;63:44-48.
- 108.** Black DW, Shaw M, Allen J. Extended release carbamazepine in the treatment of pathological gambling: an open-label study. *Prog Neuropsychopharmacol Biol Psychiatry*. 2008;32:1191-1194.
- 109.** Black DW, Shaw M, Forbush KT, Allen J. An open-label study of escitalopram in the treatment of pathological gambling. *Clin Neuropharmacol*. 2007;30:206-212.
- 110.** du Toit PL, van Kradenburg J, Niehaus D, Stein DJ. Comparison of obsessive-compulsive disorder patients with and without comorbid putative obsessive-compulsive spectrum disorders using a structured clinical interview. *Compr Psychiatry*. 2001;42:291-300.
- 111.** Hantouche EG, Lancrenon S, Bouhassira M, et al. Repeat evaluation of impulsiveness in a cohort of 155 patients with obsessive-compulsive disorder: 12 months prospective follow-up. *Encephale*. 1997;23:83-90.
- 112.** Frost RO, Meagher BM, Riskind JH. Obsessive-compulsive features in pathological lottery and scratch-ticket gamblers. *J Gambling Stud*. 2001;17:5-19.
- 113.** Forbush KT, Shaw MC, Graeber MA, et al. Neuropsychological characteristics and personality traits on pathological gambling. *CNS Spectrums*. 2008;13:306-315.
- 114.** Black DW, Goldstein RB, Noyes R, Blum N. Compulsive behaviors and obsessive-compulsive disorder (OCD): Lack of a relationship between OCD, eating disorders, and gambling. *Compr Psychiatry*. 1994;35:145-148.
- 115.** BienvENU OJ, Samuels JF, Riddle MA, et al. The relationship of obsessive-compulsive disorder to possible spectrum disorders: results from a family study. *Biol Psychiatry*. 2000;48:287-293.
- 116.** Goudriaan AE, Ossterlaan J, deBeurs E, van den Brink W. Pathological gambling: A comprehensive review of biobehavioral findings. *Neurosci Biobehav Rev*. 2004;28:123-141.
- 117.** Cavadini P, Riboldi G, Keller R, et al. Frontal lobe dysfunction in pathological gambling patients. *Biol Psychiatry*. 2002;51:334-341.
- 118.** Menzies L, Chamberlain SR, Laird AR, et al. Integrating evidence from neuroimaging and neuropsychological studies of obsessive-compulsive disorder: the orbitofrontal-striatal model revisited. *Neurosci Biobehav Rev*. 2008;525-549.
- 119.** Cunningham-Williams RM, Gattis MN, Dore PM, et al. Towards DSM-V: considering other withdrawal-like symptoms of pathological gambling disorder. *Int J Methods Psychiatr Res*. 2009;18:13-22.
- 120.** Black DW, Gaffney GR. Subclinical obsessive-compulsive disorder in children and adolescents: additional results from a "high-risk" study. *CNS Spectrums*. 2008;9(suppl 14):54-61.

NEUROETHICS OF NEUROMARKETING



While consumer neuroscience and neuromarketing has many proponents, others are also concerned that the use of neuroscience in business will allow unprecedented access to consumers' minds, opening the floodgates to pure manipulation of consumers.

As always, there are two sides of a coin. Neuromarketing both has its promises and uses, but also its perils and problems. To get to a valid use of neuroscience in marketing and consumer insights, we need to face these challenges and accommodate the practices accordingly. While academic researchers are (or should be) well versed in ethics codes of conduct, this is often not the case for commercial uses of neuroscience.

But as with most technologies, it is not the technology but the use of it that can challenge ethical uses. Two initial reservations can be stated for the direct link between unethical conduct and neuromarketing:

- 1) Companies are very aware of ethical issues, and the aspect of Corporate Social Responsibility is an important aspect that makes many companies put strong demands on the actual conduct of neuromarketing or any other marketing study
- 2) Neuromarketing, despite it's many advantages, does not in any way provide a direct link to customers' "buy button". While it is an important leg to stand on for customer insights, in no way does it make customers like purchasing zombies.

Neuroethics of neuromarketing

Emily R. Murphy^{1,2†}, Judy Illes^{1‡} and Peter B. Reiner^{1*}

¹National Core for Neuroethics, University of British Columbia, Vancouver, BC, Canada

²Center for Law and the Biosciences, Stanford Law School, Stanford, CA, USA

- *Neuromarketing is upon us. Companies are springing up to offer their clients brain-based information about consumer preferences, purporting to bypass focus groups and other marketing research techniques on the premise that directly peering into a consumer's brain while viewing products or brands is a much better predictor of consumer behavior. These technologies raise a range of ethical issues, which fall into two major categories: (1) protection of various parties who may be harmed or exploited by the research, marketing, and deployment of neuromarketing and (2) protection of consumer autonomy if neuromarketing reaches a critical level of effectiveness. The former is straightforward. The latter may or may not be problematic depending upon whether the technology can be considered to so effectively manipulate consumer behavior such that consumers are not able to be aware of the subversion. We call this phenomenon stealth neuromarketing. Academics and companies using neuromarketing techniques should adopt a code of ethics, which we propose here, to ensure beneficent and non-harmful use of the technology in consideration of both categories of ethics concerns.*

Copyright © 2008 John Wiley & Sons, Ltd.

Introduction

In 1957, the marketing executive James Vicary announced that he had increased sales of food and drink at a movie theater by secretly flashing subliminal messages with the words “Drink Coca Cola” and “Eat Popcorn”. The study was never published and may have even been a hoax (Karremans *et al.*, 2006), but the episode illustrates the public’s strong reaction to covert manipulation. An article in *The*

Nation called it “the most alarming invention since Mr. Gatling invented his gun” and *The New Yorker* stated that “minds had been entered and broken” (Moore, 1982). With growing public understanding that the brain is the mediator of behavior, the public’s reaction to neuromarketing intrusions into their brains may prove to be equally vigorous.

The term “neuromarketing” identifies a new field of research championed by both academics and self-labeled companies using advances in neuroscience that permit powerful insights into the human brain’s responses to marketing stimuli (Renvoisé and Morin, 2007; Senior *et al.*, 2007). The goals of neuromarketing studies are to obtain objective information about the inner workings of the brains of consumers without resorting to the subjective

*Correspondence to: Peter B. Reiner, Professor, National Core for Neuroethics, University of British Columbia, 2211 Wesbrook Mall, S121 Vancouver, BC V6T 2B5, Canada.

E-mail: peter.reiner@ubc.ca

†Fellow

‡Canada Research Chair in Neuroethics and Professor of Neurology

reports that have long been the mainstay of marketing studies. Thus, neuromarketing purports to provide qualitatively different information, ostensibly superior to that obtained by traditional means, about the economically valuable topic of consumer preferences.

There is, of course, nothing inherently problematic about the use of scientific technology to advance commercial interests (Eaton and Illes, 2007). But the use of technology that probes the inner workings of the human brain, especially beyond what one might divulge in traditional behavioral testing, raises substantial ethical issues. These concerns fall into two major categories: (1) protection of various parties who may be harmed or exploited by neuromarketing and (2) protection of consumer autonomy. The public outcry in response to Vicary's subliminal imagery reflects a clear ethical boundary – the autonomy violation produced intrinsic discomfort with consumers having their preferences manipulated when they did not and *could not* know as much. A similar boundary can be drawn for contemporary marketing, particularly when informed by information gleaned from novel neurotechnologies. Neuroethics, in proactively dealing with ethical issues unique to knowledge about and manipulation of the human brain, is well-positioned to offer guidance for beneficent and non-harmful deployment of neuromarketing techniques.

We will first briefly review the state of the art and state of the market in neuromarketing. The second section tackles current ethical issues in neuromarketing, which apply irrespective of the technological capabilities. We focus our discussion on potential harms to research subjects, exploitation of vulnerable niche populations, and the integrity of business relationships, public trust, and consumer confidence. The third section will thoroughly explore the most substantial neuroethical concern associated with neuromarketing: the incursion on autonomy made by neuromarketing if it achieves a level of effectiveness that amounts to consumer coercion. We conclude with pragmatic recommendations: an ethical code to be adopted by the

neuromarketing industry to prevent harms and preserve business integrity and consumer trust.

The market for neuromarketing

Although the electroencephalography (EEG) has been in use for the study of marketing preferences for over 35 years (Krugman, 1971), there is little doubt that we have entered a new age of neuromarketing in which advanced technology is being used in unprecedented ways to probe consumer preferences. A raft of peer-reviewed papers and books have appeared in recent years in which positron emission tomography (PET), functional magnetic resonance imaging (fMRI) and quantitative EEG analyses have been used to assess consumer behavior (Smith *et al.*, 2002; Dickhaut *et al.*, 2003; McClure *et al.*, 2004; Mast and Zaltman, 2005; Ahlert *et al.*, 2006; Knutson *et al.*, 2007; Koenigs and Tranel, 2008; Plassmann *et al.*, 2007; Renvoisé and Morin, 2007; Schaefer and Rotte, 2007a, b). It has already been anticipated that other neuroimaging technologies such as magnetoencephalography and cortical manipulation with transcranial magnetic stimulation, as well as combinations of modalities will be adopted by "market researchers who wish to deploy a specialized neuromarketing profile" (Senior *et al.*, 2007). We refer the reader interested in details of each technology to the recent overview by Senior *et al.* (2007) as well as other papers in this Special Issue; different neuroimaging techniques have strengths and weaknesses in temporal and spatial resolution. The choice of modality by neuromarketers will no doubt be informed by *a priori* hypotheses and pilot research about relevant brain areas and activation patterns useful for predicting actual consumer behavior.

The convergence of increased power in the form of technology and advances in our understanding of cognitive function has emboldened some to make sweeping conclusions about the power of neuromarketing. Indeed, at least ten commercial enterprises have been established with the explicit

objective of using these advanced technologies to provide neuromarketing (Emsense, FKF Applied Research, Lucid Systems, Neurofocus, Neuroco, Neurosense Limited, OTOInsights, Sales Brain, Sands Research, and Thought Sciences) and at least one US patent has been issued on the topic (Zaltman and Kosslyn, 2000). Neuromarketing studies have garnered a great deal of attention from the public, with extensive coverage in both the mainstream press (Kelly, 2002; Roston, 2004; Greene, 2007; Haq, 2007; Park, 2007; Saletan, 2007; Baker, 2008; Brainard, 2008) and internet weblogs (Dooley, 2007). One study has even cautioned against a "neurorealism" created by press coverage of novel technologies and their real or potential applications in society (Racine *et al.*, 2005).

Protection of vulnerable parties in research, selling, and representation of neuromarketing

We first consider a set of issues that merit ethical analysis irrespective of whether the most speculative claims of neuromarketing hold up to rigorous scientific analysis. Ethical development of neuromarketing requires protection of the research subjects, responsible business-to-business advertising, and accurate representation of the state of the art of the technology to the public. Each of these duties can be ensconced in an industry-wide code of ethics that we propose be adopted by all researchers and vendors of neuromarketing and enforced by a discerning marketplace of neuromarketing consumers doing business with companies voluntarily adhering to the code of ethics. Not only would adoption of a code of ethics be justified on moral grounds, but it would also serve to insulate this young and dynamic industry from accusations of irresponsible behavior.

Human subjects' protection

It is well established that federally funded scientists working in academic, government,

and commercial settings have both ethical and legal responsibilities to obtain informed consent and protect the privacy of research subjects whose brain function is probed with imaging technologies, as per the Common Rule (DHHS, 1991). The legal framework for such privacy protection in the United States is covered under PRIVACY RULE of the Department of Health and Human Services (DHHS, 2005); while this applies in some instances, it is notable that such protections are apparently absent when the subject is participating in a study being carried out for marketing purposes (Tovino, 2005). Thus standards for protecting the privacy of individuals participating in neuromarketing studies in the United States are at the very least considerably compromised. In academic and medical research centers, subjects volunteering to participate in neuroimaging-based studies are protected by Institutional Review Board guidelines, which can include strict experimental guidelines because most imaging technologies are considered to be FDA-regulated medical devices (FDA, 1998). However, when moved into commercialized and private enterprise, such subject protections may not be present, and the particularly loose restrictions surrounding studies for marketing purposes are especially worrying. Moreover, if new technologies are developed that fall outside the purview of regulatory authorities, even these protections may be lost. A key initiative for neuroethics in neuromarketing is to develop published codes of subject protections equal to those required by academic and medical research centers. fMRI, the most prevalent of the neuromarketing imaging modalities, is arguably a low-risk technology; nonetheless, risks of various sorts are inherent in all brain-imaging research and all subjects regardless of the purpose of the study are deserving of adequate protection and appropriate informed consent procedures.

Even thornier than the issue of subject protection is the notion that advanced technology in the neurosciences, in particular fMRI, might allow invasion of the inner sanctum of private thought. It bears repeating

that such a breach is not possible today and may or may not be technically achieved in the future. Nonetheless, a vigorous discussion has emerged regarding this possibility (Kulynych, 2002; Illes and Racine, 2005; Tovino, 2005; Greely, 2006; Alpert, 2007; Appelbaum, 2007; Illes, 2007; Tovino, 2007) and it is certainly worth considering how society might manage such information if it became technically feasible.

Preventing exploitation of niche populations

Special ethics review should be a minimum standard for neuromarketing research that either involves or targets vulnerable populations. Among the individuals that would fall under this umbrella are persons (or family members of persons) with neurological disease or psychological disorders, children, and other members of legally protected groups (Civil Right Act (2008) 42 U.S.C.A. §§ 2000e et seq.; Coenen, 2007). While such "segmentation" and "target marketing" is standard business practice in the marketing industry, there is a fine line between target marketing and exploitation (Sims, 1997). Our neuroethical concern is about potential harms to vulnerable persons as: (1) subjects in unregulated neuromarketing research (introduced above), (2) targeted populations who may be especially sensitive to trumped-up claims of product effectiveness based on information derived from advanced neuroscience technologies, and (3) people particularly exposed to "stealth neuromarketing" techniques that such research and development may produce (see below). Fortunately, the prevention of such harms to vulnerable persons aligns with profit motive, for as Sims (1997) points out, when targeting a particular market "maligns those it tries to serve," it undercuts its own business interests.

Responsible business-to-business advertising and public representation

It is perhaps not surprising that neuromarketing oversells its wares – an occupational hazard of

sorts. Independent critics have openly and quite rightly condemned neuromarketing efforts that overstate the benefits of the approach. The editors of the high-impact journal *Nature Neuroscience* succinctly reviewed the dangers of over interpretation of neuromarketing results, noting that the traditional skeptical approach of scientific inquiry is being displaced by a wave of media hype which suggests that fMRI "is on the verge of creating advertising campaigns that we will be unable to resist" (Editorial, 2004). In this sense, neuromarketing represents just one example of a more general problem in neuroimaging research – the question of the degree to which results which are certainly fascinating and worthy of continued attention can be used to derive *bona fide* insights about the working of the human brain (Illes and Racine, 2005), and, with particular relevance to the claims of neuromarketing, the accurate prediction of human behavior. Business consumers of neuromarketing may find their advertising dollars misspent if the technology does not live up to its claims and pass the real-world test of accurate prediction of actual consumer behavior. However, we are not overly concerned here with business-to-business relationships; harms there are primarily the potential for financial loss, but are not inherently unethical (though they may not meet industry standards of "truth-in-advertising") (Frazier, 2007). However, potential for actual harm exists if such "neurohype" around the perceived capabilities of neuromarketing create fear, anxiety, or mistrust in the general public.

Scientists working in the field of neurobiology recognize the considerable challenge involved in the translation of the brain's extraordinary connectivity – the human brain is arguably the most complex biological organ in the known universe, with tens of billions of cells, each of which make thousands of connections with other cells (Purves *et al.*, 2008) – into the complex repertoire of behavior exhibited by humans. At the same time, the general public finds color-coded images of brains in action accompanied by neuroscientific explanations to be particularly persuasive (Dumit, 2003; McCabe and Castel,

2007; Weisberg *et al.*, 2008). This tension leads to a situation where highly sophisticated scientists, subject to both public adulation and profit motive, are tempted to provide simplistic answers to what in reality are highly nuanced questions.

The neuromarketing field has already seen egregious abuse of such information. In an op-ed piece in the *New York Times*, a group of academics and neuromarketers presented a small body of unpublished data on the results of an fMRI study of political preferences of so-called 'swing voters' (Iacoboni *et al.*, 2007). The results were presented essentially as *de facto* probes into the minds of their subjects in one of the most widely read newspapers in the world; for several days after its publication, the article topped the rankings of those most frequently emailed by readers of the online version of the *New York Times*. Given widespread concerns about over-interpretation of fMRI data (Illes *et al.*, 2006b), it was notable that the op-ed piece contained none of the qualifications that would normally accompany a scholarly article in a peer-reviewed journal. Academic colleagues responded with considerable outrage in letters to the editor (Aron *et al.*, 2007) and most visibly in a scathing editorial in *Nature* (Editorial, 2007). Incidents such as this draw attention to the absence of a code of ethics for responsible media - if not academic - representation in the field of neuromarketing. Such misrepresentation can do considerable damage to the public trust of science and may even generalize in public perception to create anxieties about the perceived motivations of neuroscientists conducting human neuroimaging research. In the current climate of tightening public funds for basic research, any such anxiety threatens the future of the field, with potential harms to public health. Academic and private sectors of neuroscience research need to maintain close partnerships and work together to promote public trust and investment in neuroscience research. That trust can be earned with forthright communication and full disclosure of risks, benefits, and limitations of research findings.

Stealth neuromarketing

In our view, the most vexing of the issues raised by neuromarketing is in the realm of autonomy. One could argue that the essential objective of marketing as a discipline is to manipulate consumer behavior - effectively, a "soft" attack on autonomy. Moreover, many of the traditional tools of marketing such as focus groups and polls rely upon nuanced interpretations of human psychology to draw conclusions about consumer behavior and then use that information to inform marketing decisions. The implicit question in the present discussion is whether the new tools of neuromarketing will provide sufficient insight into human neural function to allow manipulation of the brain such that the consumer *cannot* detect the subterfuge and that such manipulations result in the desired behavior in at least some exposed persons. Such *stealth neuromarketing* is not possible with current technology, but if developed would represent a major incursion on individual autonomy. In this analysis, we deliberately consider a set of issues that will only arise with developments in technology that are yet to be realized and may never come to fruition. Nonetheless, it is in the best interest of all parties involved in the discussion that these issues are considered today rather than at some unspecified time-point in the future, possibly in response to an adverse event. To appreciate how stealth neuromarketing may come to pass, we present a short discourse on phenomena in which decision-making and motivation occur without explicit conscious awareness.

It is well established that cognition is not a monolithic process but rather one with various submodalities carrying out a variety of functions, some of which have been reasonably well delineated. Of relevance to the present discussion are two well-studied phenomena in cognitive neuroscience. The first is blindsight (Weiskrantz, 1990), in which individuals with damage to portions of their visual cortex declare themselves unable to see objects placed in the damaged portion of their visual field, but when asked to guess are easily able to

identify the object. These experiments unequivocally demonstrate that there is a distinction between perception and conscious awareness. The second phenomenon is implicit learning (Reber, 1993), in which normal subjects are presented with seemingly random strings of letters and asked to memorize them. Unknown to the individuals is the fact that there is a set of rules being employed, but their attention is directed towards memorization in the first part of the experiment. Later they are asked to describe the rules, and after protesting that they did not know there were any rules, they are asked to guess. Remarkably, subjects correctly identify the rules over 70% of the time. A recent brief report in *Science* (Aarts *et al.*, 2008) takes these phenomena one step further and demonstrates how subliminal priming effects can actually motivate and mobilize people to respond more quickly and spend extra effort on a simple motor task. The simple but elegant study provides evidence for the "human capacity to rely on mental processes in preparing and motivating behavior outside of awareness" (Aarts *et al.*, 2008). As the authors note, such responses are of considerable utility insofar as they prepare individuals to react quickly as circumstances necessitate; at the same time, subliminal priming represents an additional step towards realization of stealth neuromarketing.

If it is possible to carry out highly sophisticated cognitive tasks such as visual perception or understanding grammatical rules, as well as to enhance motivation and mobilization of voluntary action without the relevant neural computation arising to the level of perceptual awareness, then it follows that at some point in the future insights from advanced technology in the neurosciences might allow corporations, governments and others to influence decisions and actions regarding brand preference without the individual being aware of the subterfuge. We would propose such an eventuality as the sensible point at which the erosion of personal autonomy becomes so substantial that one might consider regulatory control, voluntary or otherwise, to protect the citizenry from unwanted intrusions on individual agency.

Recommendations

We conclude with a preliminary version of a code of ethics that we recommend be adopted by the neuromarketing industry. The overarching goal of this code of ethics is to promote research and development, entrepreneurship, and profitable enterprise alongside beneficent and non-harmful use of neuroimaging technology at all stages of development, deployment, and dissemination. These codes should be discussed within the neuromarketing community with the advice of independent academic researchers working in the area of neural correlates of decision-making, social behavior, and consumer preferences, as well as neuroethicists and professionals in marketing industry ethics. Proactive development of such guidelines within the professional community will provide credibility and garner greater acceptance than those that may be imposed upon the field by regulatory bodies, especially if they arise in response to adverse events (Illes *et al.*, 2003). Timeliness in this effort is critical given the rapid pace of advancements in the field.

- **Protection of research subjects.** Policies for responsibly managing clinical findings, including provision of sufficient subject protections, procedures for informed consent, and explicit protocols for dealing with incidental findings (Illes *et al.*, 2006a) are a requirement for any entity involved in brain research. Furthermore, private companies offering financial incentives for participation in research studies significantly greater than those offered in academic settings should be cautious of undue influence of such incentives, which may cross over into indirect coercion. While most technologies used by neuromarketing may be considered minimal risk, subjects should be advised and reminded of their right to withdraw from any study for any reason, including minor discomfort.
- **Protection of vulnerable niche populations from marketing exploitation.** Policies for research subjects' protection should include

additional ethics review for research done on protected or potentially vulnerable subject populations. In addition, neuromarketing-influenced advertising targeted at specific protected consumer groups should aim to beneficially serve the special needs of the population without marginalizing, maligning, or otherwise causing harm, whether psychosocial or financial in nature.

- **Full disclosure of goals, risks, and benefits.** Disclosure can be achieved through the publication of ethics principles that have been adopted to protect the privacy and autonomy of human subjects and consumers. Publication infers all aspects of the process from consent documents to reporting and advertising and applies to both written and verbal communication.
- **Accurate media and marketing representation.** Neuromarketing companies bear the burden of accurately representing their wares in media and business-to-business marketing materials. At a minimum, this standard encompasses full disclosure of scientific methods and measures of validity in mass media formats such as invited opinions, editorials, and news reports. Adherence to a code of responsible communication and truth-in-advertising will help maintain a positive and trusting public perception of brain science research as well as promote development of effective technologies.
- **Internal and external validity.** Eaton and Illes (2007) have outlined the challenges in initial and sustained product validity in the commercialization of any neurotechnology. We extend their recommendations here to any marketing product influenced by neuromarketing research with particular attention to the point that the validity questions "arise most acutely for neurotechnology that can be deployed without a regulatory gatekeeper, such as the FDA" (Eaton and Illes, 2007). At a minimum, internal validity checks should ensure a sufficiently comprehensive research database to provide meaningful and effective results to neuromarketing consumers. External and

sustained validity will require neuromarketers to align their product with changing technologies and expanding neuroscience knowledge. Maintenance of safety and efficacy verification in any research, development, and deployment of neuromarketing is absolutely required.

These recommendations form the basis for immediate and short-term action in the neuromarketing community and longer-term empirical research. Multidisciplinary collaboration will enable efficient and positive progress along this continuum.

In the 50 years since Vicary's subliminal imagery marketing stunt, interest in the possibility that neuroscience might be used in the service of a marketing agenda has remained robust, with the current resurgence of interest and proliferation of companies in the new neuromarketing being noteworthy. It should be emphasized that there is no evidence that at the present time that any advanced neurotechnology permits the types of insights and subsequent manipulation that Vicary envisaged (Illes and Racine, 2005; Illes, 2007). However, the fact that one must insert qualifiers such as "at the present time" provides ample reason to carefully consider the implications that such a development might have and the means by which it might be sensibly managed. In the meantime, there are a host of ethical issues in the research, marketing, and deployment of neuromarketing on the table right now. Such proactive conduct is at the heart of the neuroethical agenda.

Acknowledgements

This study was supported by grants from the Canadian Institutes of Health Research (Institute for Neuroscience, Mental Health and Addiction and the Ethics Office of CIHR), the Canadian Foundation for Innovation, the British Columbia Knowledge Development Fund, the Dana Foundation, the Greenwall Foundation, and the MacArthur Foundation-funded Law and Neuroscience Project.

Biographical notes

Emily R. Murphy is a research fellow at the Stanford Law School Center for Law and the Biosciences, a research fellow on the MacArthur Foundation-funded Law and Neuroscience Project, and a visiting fellow at the National Core for Neuroethics, UBC. Dr Murphy received her Ph.D. in behavioral neuroscience from the University of Cambridge, where her work focused on neural and neurochemical correlates of behavioral inhibition and behavioral flexibility. Her current research interests are at the intersection of law and neuroscience, particularly in the areas of neuroimaging as evidence, and the neuroscience of decision-making and drug addiction.

Judy Illes is the Director of the National Core for Neuroethics at UBC and a pioneer in the field with deep interests in ethical, social, and policy challenges at the intersection of the neurosciences and biomedical ethics. Dr Illes has written numerous books, edited volumes and articles. Her latest book, *Neuroethics: Defining the Issues in Theory, Practice and Policy*, was published by Oxford University Press in 2006. Dr Illes is a member of the Internal Advisory Board for the Institute of Neurosciences, Mental Health and Addiction (INMHA) of the Canadian Institutes of Health Research (CIHR), the Institute of Medicine, Forum on Neuroscience on Neurological Disorders, the Dana Alliance for Brain Initiatives, and co-Chair of the Committee on Women in Neuroscience for the Society for Neuroscience.

Peter B. Reiner is a Professor in the National Core for Neuroethics at the University of British Columbia. Following a distinguished career as a research scientist studying the neurobiology of behavioral states and the molecular underpinnings of neurodegenerative disease, Dr Reiner became President & CEO of Active Pass Pharmaceuticals, a drug discovery company that he founded to tackle the scourge of Alzheimer's disease. Upon returning to academic life, Dr Reiner refocused his scholarly work in the area of neuroethics, focusing upon the ethical issues associated with commercia-

lization of neuroscience, with a particular interest in the emerging debate over cognitive enhancers.

References

- Aarts H, Custers R, Marien H. 2008. Preparing and motivating behavior outside of awareness. *Science* **319**: 1639.
- Ahlert D, Kenning P, Plassmann H. 2006. A window to the consumer's mind: application of functional brain imaging techniques to advertising research. In *International Advertising and Communication*. 163–178.
- Alpert S. 2007. Brain privacy: how can we protect it? *The American Journal of Bioethics* **7**: 70–73.
- Appelbaum PS. 2007. Law & psychiatry: the new lie detectors: neuroscience, deception, and the courts. *Psychiatric Services* **58**: 460–462.
- Aron A, Badre D, Brett M, Cacioppo J, Chambers C, Cools R, Engel S, D'Esposito M, Frith C, Harmon-Jones E, Jonides J, Knutson B, Phelps L, Poldrack R, Wager T, Wagner A, Winkielman P. November 14 2007. Politics and the brain (Letter to the Editor). *New York Times*.
- Baker S. 2008. What you really want to buy. *Business Week*. http://www.businessweek.com/technology/content/jan2008/tc20080127697425.htm?chan=technology_technology+index+page_top+stories [10 March 2008].
- Brainard C. 2008. Beware of "neuropunditry". *Columbia Journalism Review*. http://www.cjr.org/campaign_desk/beware_of_neuropunditry.php [11 March 2008].
- Coenen D. 2007. The future of footnote four. *Georgia Law Review* **41**: 797.
- DHHS, Department of Health and Human Services, NIH, and OPRR. 1991. 45 CFR 46.
- DHHS, Department of Health and Human Services, 45 C.F.R. Parts 160 and 164, 2005.
- Dickhaut J, McCabe K, Nagode JC, Rustichini A, Smith K, Pardo JV. 2003. The impact of the certainty context on the process of choice. *Proceedings of the National Academy of Sciences* **100**: 3536–3541.
- Dooley R. 2007. Neuromarketing blog [Internet], <http://www.neurosciencemarketing.com/blog/>, accessed on November 26, 2007.

- Dumit J. 2003. *Picturing Personhood: Brain Scans and Biomedical Identity*. Princeton University Press: Princeton, NJ.
- Eaton ML, Illes J. 2007. Commercializing cognitive neurotechnology - the ethical terrain. *Nature Biotechnology* **25**: 393-397.
- Editorial. 2004. Brain scam? *Nature Neuroscience* **7**: 683.
- Editorial. 2007. Mind games. *Nature* **450**: 457.
- FDA. 1998. Guidance for institutional review board and clinical investigators. Medical devices.
- Frazier M. 2007. Hidden persuasion or junk science? *Advertising Age* **78**: 1-2.
- Greely HT. 2006. The social effects of advances in neuroscience: legal problems, legal perspectives. In *Neuroethics: Defining the Issues in Theory, Practice, and Policy*, J Illes (ed). Oxford University Press: Oxford; 245-264.
- Greene K. 2007. Brain sensor for market research. *Technology Review*. <http://www.technologyreview.com/Biztech/19833/> [10 March 2008].
- Haq A. 2007. This is your brain on advertising. *Business Week*. http://www.businessweek.com/globalbiz/content/oct2007/gb2007108_286282.htm?chan=top+news_top+news+index_global+business [11 March 2008].
- Iacoboni M, Freedman J, Kaplan J, Jamieson KH, Freedman T, Knapp B, Fitzgerald K. 2007. This is your brain on politics. *The New York Times*. <http://www.nytimes.com/2007/11/11/opinion/11freedman.html?r=1&sq=marco%20iacoboni%202007&st=cse&oref=slogin&scp=1&pagewanted=all> [18 November 2007].
- Illes J. 2007. Empirical neuroethics. Can brain imaging visualize human thought? Why is neuroethics interested in such a possibility? *EMBO Reports* **8**(Spec No): S57-S60.
- Illes J, Racine E. 2005. Imaging or imagining? A neuroethics challenge informed by genetics. *The American Journal of Bioethics* **5**: 5-18.
- Illes J, Fan E, Koenig BA, Raffin TA, Kann D, Atlas SW. 2003. Self-referred whole-body CT imaging: current implications for health care consumers. *Radiology* **228**: 346-351.
- Illes J, Kirschen MP, Edwards E, Stanford LR, Bandettini P, Cho MK, Ford PJ, Glover GH, Kulynych J, Macklin R, et al. 2006a. Ethics. Incidental findings in brain imaging research. *Science* **311**: 783-784.
- Illes J, Racine E, Kirschen MP. 2006b. A picture is worth 100 words, but which 1000?. In *Neuroethics: Defining the Issues in Theory, Practice, and Policy*, J Illes (ed). Oxford University Press: Oxford; 148-168.
- Karremans JC, Stroebe W, Claus J. 2006. Beyond Vicary's fantasies: the impact of subliminal priming and brand choice. *Journal of Experimental Social Psychology* **42**: 792-798.
- Kelly M. 2002. The science of shopping. In CBC's Marketplace (Canada).
- Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G. 2007. Neural predictors of purchases. *Neuron* **53**: 147-156.
- Koenigs M, Tranel D. 2008. Prefrontal cortex damage abolishes brand-cued changes in cola preference. *Social Cognitive Affect Neuroscience* **3**: 1-6.
- Krugman HE. 1971. Brain wave measures of media involvement. *Journal of Advertising Research* **11**: 3-9.
- Kulynych J. 2002. Legal and ethical issues in neuroimaging research: human subjects protection, medical privacy, and the public communication of research results. *Brain and Cognition* **50**: 345-357.
- Mast FW, Zaltman G. 2005. A behavioral window on the mind of the market: an application of the response time paradigm. *Brain Research Bulletin* **67**: 422-427.
- McCabe DP, Castel AD. 2007. Seeing is believing: the effect of brain images on judgments of scientific reasoning. *Cognition* **107**: 343-352.
- McClure SM, Li J, Tomlin D, Cypert KS, Montague LM, Montague PR. 2004. Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* **44**: 379-387.
- Moore TE. 1982. Subliminal advertising: what you see is what you get. *Journal of Marketing* **46**: 38-47.
- Park A. 2007. Marketing to your mind. JT Time. <http://www.time.com/time/magazine/article/0,9171,1580370,00.html> [11 March 2008].
- Plassmann H, Ambler T, Braeutigam S, Kenning P. 2007. What can advertisers learn from neuroscience? *International Journal of Advertising* **26**: 151-175.
- Purves D, Fitzpatrick D, Augustine GJ, Katz LC. 2008. *Neuroscience*, 4th edn. Sinauer: Sunderland, MA.

- Racine E, Bar-Ilan O, Illes J. 2005. fMRI in the public eye. *Nature Reviews Neuroscience* **6**: 159-164.
- Reber AS. 1993. *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford University Press: New York.
- Renvoisé P, Morin C. 2007. *Neuromarketing: Understanding the "Buy Button" in Your Customer's Brain*. T. Nelson: Nashville, TN.
- Roston E. 2004. The why of buy. *Time*. <http://www.time.com/time/insidebiz/article/0,9171,1101040308-596161,00.html>.
- Saletan W. 2007. Peering into the soul. *Washington Post*. <http://www.washingtonpost.com/wpdyn/content/article/2007/03/16/AR2007031602672.html> [11 March 2008].
- Schaefer M, Rotte M. 2007a. Thinking on luxury or pragmatic brand products: brain responses to different categories of culturally based brands. *Brain Research* **1165**: 98-104.
- Schaefer M, Rotte M. 2007b. Favorite brands as cultural objects modulate reward circuit. *Neuroreport* **18**: 141-145.
- Senior C, Smythe H, Cooke R, Shaw RL, Peel E. 2007. Mapping the mind for the modern market researcher. *Qualitative Market Research* **10**: 153-167.
- Sims R. 1997. When does target marketing become exploitation? *Marketing News* **31**: 10.
- Smith K, Dickhaut J, McCabe K, Pardo JV. 2002. Neuronal substrates for choice under ambiguity, risk, gains, and losses. *Management Science* **48**: 711-718.
- Tovino SA. 2005. The confidentiality and privacy implications of functional magnetic resonance imaging. *The Journal of Law, Medicine & Ethics* **33**: 844-850.
- Tovino SA. 2007. Functional neuroimaging and the law: trends and directions for future scholarship. *The American Journal of Bioethics* **7**: 44-56.
- Weisberg DS, Keil FC, Goodstein J, Rawson E, Gray JR. 2008. The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience* **20**: 470-477.
- Weiskrantz L. 1990. The Ferrier lecture, 1989. Outlooks for blindsight: explicit methodologies for implicit processes. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)* **239**: 247-278.
- Zaltman G, Kosslyn SM. 2000. Neuroimaging as a marketing tool. USPTO, ed. (United States).



NIH Public Access

Author Manuscript

J Cogn Neurosci. Author manuscript; available in PMC 2009 November 18.

Published in final edited form as:

J Cogn Neurosci. 2008 March ; 20(3): 470–477. doi:10.1162/jocn.2008.20040.

The Seductive Allure of Neuroscience Explanations

Deena Skolnick Weisberg, Frank C. Keil, Joshua Goodstein, Elizabeth Rawson, and Jeremy R. Gray
Yale University

Abstract

Explanations of psychological phenomena seem to generate more public interest when they contain neuroscientific information. Even irrelevant neuroscience information in an explanation of a psychological phenomenon may interfere with people's abilities to critically consider the underlying logic of this explanation. We tested this hypothesis by giving naïve adults, students in a neuroscience course, and neuroscience experts brief descriptions of psychological phenomena followed by one of four types of explanation, according to a 2 (good explanation vs. bad explanation) \times 2 (without neuroscience vs. with neuroscience) design. Crucially, the neuroscience information was irrelevant to the logic of the explanation, as confirmed by the expert subjects. Subjects in all three groups judged good explanations as more satisfying than bad ones. But subjects in the two nonexpert groups additionally judged that explanations with logically irrelevant neuroscience information were more satisfying than explanations without. The neuroscience information had a particularly striking effect on nonexperts' judgments of bad explanations, masking otherwise salient problems in these explanations.

INTRODUCTION

Although it is hardly mysterious that members of the public should find psychological research fascinating, this fascination seems particularly acute for findings that were obtained using a neuropsychological measure. Indeed, one can hardly open a newspaper's science section without seeing a report on a neuroscience discovery or on a new application of neuroscience findings to economics, politics, or law. Research on nonneural cognitive psychology does not seem to pique the public's interest in the same way, even though the two fields are concerned with similar questions.

The current study investigates one possible reason why members of the public find cognitive neuroscience so particularly alluring. To do so, we rely on one of the functions of neuroscience information in the field of psychology: providing explanations. Because articles in both the popular press and scientific journals often focus on how neuroscientific findings can help to explain human behavior, people's fascination with cognitive neuroscience can be redescribed as people's fascination with *explanations* involving a neuropsychological component.

However, previous research has shown that people have difficulty reasoning about explanations (for reviews, see Keil, 2006; Lombrozo, 2006). For instance, people can be swayed by teleological explanations when these are not warranted, as in cases where a nonteleological process, such as natural selection or erosion, is actually implicated (Lombrozo & Carey, 2006; Kelemen, 1999). People also tend to rate longer explanations as more similar

to experts' explanations (Kikas, 2003), fail to recognize circularity (Rips, 2002), and are quite unaware of the limits of their own abilities to explain a variety of phenomena (Rozenblit & Keil, 2002). In general, people often believe explanations because they find them intuitively satisfying, not because they are accurate (Trout, 2002).

In line with this body of research, we propose that people often find neuroscience information alluring because it interferes with their abilities to judge the quality of the psychological explanations that contain this information. The presence of neuroscience information may be seen as a strong marker of a good explanation, regardless of the actual status of that information within the explanation. That is, something about seeing neuroscience information may encourage people to believe they have received a scientific explanation when they have not. People may therefore uncritically accept any explanation containing neuroscience information, even in cases when the neuroscience information is irrelevant to the logic of the explanation.

To test this hypothesis, we examined people's judgments of explanations that either do or do not contain neuroscience information, but that otherwise do not differ in content or logic. All three studies reported here used a 2 (explanation type: good vs. bad) \times 2 (neuroscience: without vs. with) design. This allowed us to see both people's baseline abilities to distinguish good psychological explanations from bad psychological explanations as well as any influence of neuroscience information on this ability. If logically irrelevant neuroscience information affects people's judgments of explanations, this would suggest that people's fascination with neuropsychological explanations may stem from an inability or unwillingness to critically consider the role that neuroscience information plays in these explanations.

EXPERIMENT 1

Methods

Subjects—There were 81 participants in the study (42 women, 37 men, 2 unreported; mean age = 20.1 years, $SD = 4.2$ years, range = 18–48 years, based on 71 reported ages). We randomly assigned 40 subjects to the Without Neuroscience condition and 41 to the With Neuroscience condition. Subjects thus saw explanations that either always did or always did not contain neuroscience information. We used this between-subjects design to prevent subjects from directly comparing explanations that did and did not contain neuroscience, providing a stronger test of our hypothesis.

Materials—We wrote descriptions of 18 psychological phenomena (e.g., mutual exclusivity, attentional blink) that were meant to be accessible to a reader untrained in psychology or neuroscience. For each of these items, we created two types of explanations, good and bad, neither of which contained neuroscience. The good explanations in most cases were the genuine explanations that the researchers gave for each phenomenon. The bad explanations were circular restatements of the phenomenon, hence, not explanatory (see Table 1 for a sample item).

For the With Neuroscience conditions, we added neuroscience information to the good and bad explanations from the Without Neuroscience conditions. The added neuroscience information had three important features: (1) It always specified that the area of activation seen in the study was an area already known to be involved in tasks of this type, circumventing the interpretation that the neuroscience information added value to the explanation by localizing the phenomenon. (2) It was always identical or nearly identical in the good explanation and the bad explanation for a given phenomenon. Any general effect of neuroscience information on judgment should thus be seen equally for good explanations and bad explanations. Additionally, any differences that may occur between the good explanation and bad explanation conditions would be highly unlikely to be due to any details of the neuroscience

information itself. (3) Most importantly, in no case did the neuroscience information alter the underlying logic of the explanation itself. This allowed us to test the effect of neuroscience information on the task of evaluating explanations, independent of any value added by such information.¹ Before the study began, three experienced cognitive neuroscientists confirmed that the neuroscience information did not add value to the explanations.

Procedure—Subjects were told that they would be rating explanations of scientific phenomena, that the studies they would read about were considered solid, replicable research, and that the explanations they would read were not necessarily the real explanations for the phenomena. For each of the 18 stimuli, subjects read a one-paragraph description of the phenomenon followed by an explanation of that phenomenon. They rated how satisfying they found the explanation on a 7-point scale from -3 (*very unsatisfying*) to $+3$ (*very satisfying*) with 0 as the neutral midpoint.

Results and Discussion

Preliminary analyses revealed no differences in performance based on sex or level of education, so ratings were collapsed across these variables for the analyses. Additionally, subjects tended to respond similarly to all 18 items (Cronbach's $\alpha = .79$); the set of items had acceptable psychometric reliability as a measure of the construct of interest.

Our primary goal in this study was to discover what effect, if any, the addition of neuroscience information would have on subjects' ratings of how satisfying they found good and bad explanations. We analyzed the ratings using a 2 (good explanation vs. bad explanation) \times 2 (without neuroscience vs. with neuroscience) repeated measures analysis of variance (ANOVA; see Figure 1).

There was a significant main effect of explanation type [$F(1, 79) = 144.8, p < .01$], showing that good explanations ($M = 0.88, SE = 0.10$) are rated as significantly more satisfying than bad explanations ($M = -0.28, SE = 0.12$). That is, subjects were accurate in their assessments of explanations in general, finding good explanations to be better than bad ones.

There was also a significant main effect of neuroscience [$F(1, 79) = 6.5, p < .05$]. Explanations with neuroscience information ($M = 0.53, SE = 0.13$) were rated as significantly more satisfying than explanations that did not include neuroscience information ($M = 0.06, SE = 0.13$). Adding irrelevant neuroscience information thus somehow impairs people's baseline ability to make judgments about explanations.

We also found a significant interaction between explanation type and neuroscience information [$F(1, 79) = 18.8, p < .01$]. Post hoc tests revealed that although the ratings for good explanations were not different without neuroscience ($M = 0.86, SE = 0.11$) than with neuroscience ($M = 0.90, SE = 0.16$), ratings for bad explanations were significantly lower for explanations without neuroscience ($M = -0.73, SE = 0.14$) than explanations with neuroscience ($M = 0.16, SE = 0.16$). Note that this difference is not due to a ceiling effect; ratings of good explanations are still significantly below the top of the scale [$t(80) = -21.38, p < .01$]. This interaction indicates that it is not the mere presence of verbiage about neuroscience that encourages people to think more favorably of an explanation. Rather, neuroscience information seems to have the specific

¹Because we constructed the stimuli in the With neuroscience conditions by modifying the explanations from the Without Neuroscience conditions, both the good and the bad explanations in the With Neuroscience conditions appear less elegant and less parsimonious than their without-neuroscience counterparts, as can be seen in Table 1. But this design provides an especially stringent test of our hypothesis: We expect that explanations with neuroscience will be judged as more satisfying than explanations without, despite cosmetic and logical flaws.

effect of making bad explanations look significantly more satisfying than they would without neuroscience.

This puzzling differential effect of neuroscience information on the bad explanations may occur because participants gave the explanations a more generous interpretation than we had expected. Our instructions encouraged participants to think of the explanations as being provided by knowledgeable researchers, so they may have considered the explanations less critically than we would have liked. If participants were using somewhat relaxed standards of judgment, then a group of subjects that is specifically trained to be more critical of judging explanations should not fall prey to the effect of added neuroscience information, or at least not as much.

Experiment 2 addresses this issue by testing a group of subjects trained to be critical in their judgments: students in an intermediate-level cognitive neuroscience class. These students were receiving instruction on the basic logic of cognitive neuroscience experiments and on the types of information that are relevant to drawing conclusions from neuroscience studies. We predicted that this instruction, together with their classroom experience of carefully analyzing neuroscience experiments, would eliminate or dampen the impact of the extraneous neuroscience information.

EXPERIMENT 2

Methods

Subjects and Procedure—Twenty-two students (10 women; mean age = 20.7 years, $SD = 2.6$ years, range = 18–31 years) were recruited from an introductory cognitive neuroscience class and received no compensation for their participation. They were informed that although participation was required for the course, the results of the experiment would have no impact on their class performance and would not be known by their professor until after their grades had been posted. They were additionally allowed to choose whether their data could be used in the published research study, and all students elected to have their data included.

Subjects were tested both at the beginning of the semester and at the end of the semester, prior to the final exam.

The stimuli and task were identical to Experiment 1, with one exception: Both main variables of explanation type and presence of neuroscience were within-subject due to the small number of participants.

Results and Discussion

Preliminary analyses showed no differences in performance based on class year, so this variable is not considered in the main analyses. There was one significant interaction with sex that is discussed shortly. Responses to the items were again acceptably consistent (Cronbach's $\alpha = .74$).

As with the novices in Experiment 1, we tested whether the addition of neuroscience information affects judgments of good and bad explanations. For the students in this study, we additionally tested the effect of training on evaluations of neuroscience explanations. We thus analyzed the students' ratings of explanatory satisfaction using a 2 (good explanation vs. bad explanation) \times 2 (without neuroscience vs. with neuroscience) \times 2 (preclass test vs. postclass test) repeated measures ANOVA (see Figure 2).

We found a significant main effect of explanation type [$F(1, 21) = 50.9, p < .01$], confirming that the students judged good explanations ($M = 0.37, SE = 0.14$) to be more satisfying than bad explanations ($M = -0.43, SE = 0.19$).

Although Experiment 1 found a strong effect of the presence of neuroscience information in explanations, we had hypothesized that students in a neuroscience course, who were learning to be critical consumers of neuroscience information, would not show this effect. However, the data failed to confirm this hypothesis; there was a significant main effect of neuroscience [$F(1, 21) = 47.1, p < .01$]. Students, like novices, judged that explanations with neuroscience information ($M = 0.43, SE = 0.17$) were more satisfying than those without neuroscience information ($M = -0.49, SE = 0.16$).

There was additionally an interaction effect between explanation type and presence of neuroscience [$F(1, 21) = 8.7, p < .01$], as in Experiment 1. Post hoc analyses indicate that this interaction happens for the same reason as in Experiment 1: Ratings of bad explanations increased reliably more with the addition of neuroscience than did good explanations. Unlike the novices, the students judged that both good explanations and bad explanations were significantly more satisfying when they contained neuroscience, but the bad explanations were judged to have improved more dramatically, based on a comparison of the differences in ratings between explanations with and without neuroscience [$t(21) = 2.98, p < .01$]. Specialized training thus did not discourage the students from judging that irrelevant neuroscience information somehow contributed positively to both types of explanation.

Additionally, our analyses found no main effect of time, showing that classroom training did not affect the students' performance. Ratings before taking the class and after completing the class were not significantly different [$F(1, 21) = 0.13, p > .10$], and there were no interactions between time and explanation type [$F(1, 21) = 0.75, p > .10$] or between time and presence of neuroscience [$F(1, 21) = 0.0, p > .10$], and there was no three-way interaction among these variables [$F(1, 21) = 0.31, p > .10$]. The only difference between the preclass data and the postclass data was a significant interaction between sex and neuroscience information in the pre-class data [$F(1, 20) = 8.5, p < .01$], such that the difference between women's preclass satisfaction ratings for the Without Neuroscience and the With Neuroscience conditions was significantly larger than this difference in the men's ratings. This effect did not hold in the postclass test, however. These analyses strongly indicate that whatever training subjects received in the neuroscience class did not affect their performance in the task.

These two studies indicate that logically irrelevant neuroscience information has a reliably positive impact on both novices' and students' ratings of explanations, particularly bad explanations, that contain this information. One concern with this conclusion is our assumption that the added neuroscience information really was irrelevant to the explanation. Although we had checked our items with cognitive neuroscientists beforehand, it is still possible that subjects interpreted some aspect of the neuroscience information as logically relevant or content-rich, which would justify their giving higher ratings to the items with neuroscience information. The subjects' differential performance with good and bad explanations speaks against this interpretation, but perhaps something about the neuroscience information genuinely did improve the bad explanations.

Experiment 3 thus tests experts in neuroscience, who would presumably be able to tell if adding neuroscience information should indeed make these explanations more satisfying. Are experts immune to the effects of neuroscience information because their expertise makes them more accurate judges? Or are experts also somewhat seduced by the allure of neuroscience information?

EXPERIMENT 3

Methods

Subjects and Procedure—Forty-eight neuroscience experts participated in the study (29 women, 19 men; mean age = 27.5 years, $SD = 5.3$ years, range = 21–45 years). There were 28 subjects in the Without Neuroscience condition and 20 subjects in the With Neuroscience condition.

We defined our expert population as individuals who are about to pursue, are currently pursuing, or have completed advanced degrees in cognitive neuroscience, cognitive psychology, or strongly related fields. Our participant group contained 6 participants who had completed college, 29 who were currently in graduate school, and 13 who had completed graduate school.

The materials and procedure in this experiment were identical to Experiment 1, with the addition of four demographic questions in order to confirm the expertise of our subjects. We asked whether they had ever participated in a neuroscience study, designed a neuroscience study, designed a psychological study that did not necessarily include a neuroscience component, and studied neuroscience formally as part of a course or lab group. The average score on these four items was 2.9 ($SD = 0.9$), indicating a high level of expertise among our participants.

Results and Discussion

Preliminary analyses revealed no differences in performance based on sex or level of education, so all subsequent analyses do not consider these variables. We additionally found acceptably consistent responding to the 18 items (Cronbach's $\alpha = .71$).

We analyzed subjects' ratings of explanatory satisfaction in a 2 (good explanation vs. bad explanation) \times 2 (without neuroscience vs. with neuroscience) repeated measures ANOVA (see Figure 3).

We found a main effect of explanation type [$F(1, 46) = 54.9, p < .01$]. Just like the novices and students, the experts rated good explanations ($M = 0.19, SE = 0.11$) as significantly more satisfying than bad ones ($M = 0.99, SE = 0.14$).

Unlike the data from the other two groups, the experts' data showed no main effect of neuroscience, indicating that subjects rated explanations in the same way regardless of the presence of neuroscience information [$F(1, 46) = 1.3, p > .10$].

This lack of a main effect must be interpreted in light of a significant interaction between explanation type and presence of neuroscience [$F(1, 46) = 8.9, p < .01$]. Post hoc analyses reveal that this interaction is due to a differential effect of neuroscience on the good explanations: Good explanations with neuroscience ($M = -0.22, SE = 0.21$) were rated as significantly less satisfying than good explanations without neuroscience [$M = 0.41, SE = 0.13$; $F(1, 46) = 8.5, p < .01$]. There was no change in ratings for the bad explanations (without neuroscience $M = -1.07, SE = 0.19$; with neuroscience $M = -0.87, SE = 0.21$). This indicates that experts are so attuned to proper uses of neuroscience that they recognized the insufficiency of the neuroscience information in the With Neuroscience condition. This recognition likely led to the drop in satisfaction ratings for the good explanations, whereas bad explanations could not possibly have been improved by what the experts knew to be an improper application of neuroscience information. Informal post hoc questioning of several participants in this study indicated that they were indeed sensitive to the awkwardness and irrelevance of the neuroscience information in the explanations.

These results from expert subjects confirm that the neuroscience information in the With Neuroscience conditions should not be seen as adding value to the explanations. The results from the two nonexpert groups are thus due to these subjects' misinterpretations of the neuroscience information, not the information itself.

GENERAL DISCUSSION

Summary of Results

The three experiments reported here explored the impact of adding scientific-sounding but empirically and conceptually uninformative neuroscience information to both good and bad psychological explanations. Three groups of subjects (novices, neuroscience class students, and neuroscience experts) read brief descriptions of psychological phenomena followed by a good or bad explanation that did or did not contain logically irrelevant neuroscience information. Although real neuropsychological data certainly can provide insight into behavior and into psychological mechanisms, we specifically wanted to investigate the possible effects of the presence of neuroscience information, regardless of the role that this information plays in an explanation. The neuroscience information in the With Neuroscience condition thus did not affect the logic or content of the psychological explanations, allowing us to see whether the mere mention of a neural process can affect subjects' judgments of explanations.

We analyzed subjects' ratings of how satisfying they found the explanations in the four conditions. We found that subjects in all groups could tell the difference between good explanations and bad explanations, regardless of the presence of neuroscience. Reasoning about these types of explanations thus does not seem to be difficult in general because even the participants in our novice group showed a robust ability to differentiate between good and bad explanations.

Our most important finding concerns the effect that explanatorily irrelevant neuroscience information has on subject's judgments of the explanations. For novices and students, the addition of such neuroscience information encouraged them to judge the explanations more favorably, particularly the bad explanations. That is, extraneous neuroscience information makes explanations look more satisfying than they actually are, or at least more satisfying than they otherwise would be judged to be. The students in the cognitive neuroscience class showed no benefit of training, demonstrating that only a semester's worth of instruction is not enough to dispel the effect of neuroscience information on judgments of explanations. Many people thus systematically misunderstand the role that neuroscience should and should not play in psychological explanations, revealing that logically irrelevant neuroscience information can be seductive—it can have much more of an impact on participants' judgments than it ought to.

However, the impact of superfluous neuroscience information is not unlimited. Although novices and students rated bad explanations as more satisfying when they contained neuroscience information, experts did not. In fact, subjects in the expert group tended to rate good explanations with neuroscience information as worse than good explanations without neuroscience, indicating their understanding that the added neuroscience information was inappropriate for the phenomenon being described. There is thus some noticeable benefit of extended and specific training on the judgment of explanations.

Why are Nonexperts Fooled?

Nonexperts judge explanations with neuroscience information as more satisfying than explanations without neuroscience, especially bad explanations. One might be tempted to conclude from these results that neuroscience information in explanations is a powerful clue to the goodness of explanations; nonexperts who see neuroscience information automatically

judge explanations containing it more favorably. This conclusion suggests that these two groups of subjects fell prey to a reasoning heuristic (e.g., Shafir, Smith, & Osherson, 1990; Tversky & Kahneman, 1974, 1981). A plausible heuristic might state that explanations involving more technical language are better, perhaps because they look more “scientific.” The presence of such a heuristic would predict that subjects should judge all explanations containing neuroscience information as more satisfying than all explanations without neuroscience, because neuroscience is itself a cue to the goodness of an explanation.

However, this was not the case in our data. Both novices and students showed a differential impact of neuroscience information on their judgments such that the ratings for bad explanations increased much more markedly than ratings for good explanations with the addition of neuroscience information. This interaction effect suggests that an across-the-board reasoning heuristic is probably not responsible for the nonexpert subjects’ judgments.

We see a closer affinity between our work and the so-called seductive details effect (Harp & Mayer, 1998; Garner, Alexander, Gillingham, Kulikowich, & Brown, 1991; Garner, Gillingham, & White, 1989). Seductive details, related but logically irrelevant details presented as part of an argument, tend to make it more difficult for subjects to encode and later recall the main argument of a text. Subjects’ attention is diverted from important generalizations in the text toward these interesting but irrelevant details, such that they perform worse on a memory test and have a harder time extracting the most important points in the text.

Despite the strength of this seductive details effect in this previous work and in our current work, it is not immediately clear *why* nonexpert participants in our study judged that seductive details, in the form of neuroscience information, made the explanations we presented more satisfying. Future investigations into this effect could answer this question by including qualitative measures to determine precisely how subjects view the differences among the explanations. In the absence of such data, we can question whether something about neuroscience information in particular did the work of fooling our subjects. We suspect not—other kinds of information besides neuroscience could have similar effects. We focused the current experiments on neuroscience because it provides a particularly fertile testing ground, due to its current stature both in psychological research and in the popular press. However, we believe that our results are not necessarily limited to neuroscience or even to psychology. Rather, people may be responding to some more general property of the neuroscience information that encouraged them to find the explanations in the With Neuroscience condition more satisfying.

To speculate about the nature of this property, people seeking explanations may be biased to look for a simple reductionist structure. That is, people often hear explanations of “higher-level” or macroscopic phenomena that appeal to “lower-level” or microscopic phenomena. Because the neuroscience explanations in the current study shared this general format of reducing psychological phenomena to their lower-level neuroscientific counterparts, participants may have jumped to the conclusion that the neuroscience information provided them with a physical explanation for a behavioral phenomenon. The mere mention of a lower level of analysis may have made the bad behavioral explanations seem connected to a larger explanatory system, and hence more insightful. If this is the case, other types of logically irrelevant information that tap into a general reductionist framework could encourage people to judge a wide variety of poor explanations as satisfying.

There are certainly other possible mechanisms by which neuroscience information may affect judgments of explanations. For instance, neuroscience may illustrate a connection between the mind and the brain that people implicitly believe not to exist, or not to exist in such a strong way (see Bloom, 2004a). Additionally, neuroscience is associated with powerful visual

imagery, which may merely attract attention to neuroscience studies but which is also known to interfere with subjects' abilities to explain the workings of physical systems (Hayes, Huleatt, & Keil, in preparation) and to render scientific claims more convincing (McCabe & Castel, in press). Indeed, it is possible that "pictures of blobs on brains seduce one into thinking that we can now directly observe psychological processes" (Henson, 2005, p. 228). However, the mechanism by which irrelevant neuroscience information affects judgment may also be far simpler: Any meaningless terminology, not necessarily scientific jargon, can change behavior. Previous studies have found that providing subjects with "placebic" information (e.g., "May I use the Xerox machine; I have to make copies?") increases compliance with a request over and above a condition in which the researcher simply makes the request (e.g., "May I use the Xerox machine?") (Langer, Blank, & Chanowitz, 1978).

These characteristics of neuroscience information may singly or jointly explain why subjects judged explanations containing neuroscience information as generally more satisfying than those that did not. But the most important point about the current study is not that neuroscience information itself causes subjects to lose their grip on their normally well-functioning judgment processes. Rather, neuroscience information happens to represent the intersection of a variety of properties that can conspire together to impair judgment. Future research should aim to tease apart which properties are most important in this impairment, and indeed, we are planning to follow up on the current study by examining comparable effects in other special sciences. We predict that any of these properties alone would be sufficient for our effect, but that they are more powerful in combination, hence especially powerful for the case of neuroscience, which represents the intersection of all four.

Regardless of the breadth of our effect or the mechanism by which it occurs, the mere fact that irrelevant information can interfere with people's judgments of explanations has implications for how neuroscience information in particular, and scientific information in general, is viewed and used outside of the laboratory. Neuroscience research has the potential to change our views of personal responsibility, legal regulation, education, and even the nature of the self (Farah, 2005; Bloom, 2004b). To take a recent example, some legal scholars have suggested that neuroimaging technology could be used in jury selection, to ensure that jurors are free of bias, or in questioning suspects, to ensure that they are not lying (Rosen, 2007). Given the results reported here, such evidence presented in a courtroom, a classroom, or a political debate, regardless of the scientific status or relevance of this evidence, could strongly sway opinion, beyond what the evidence can support (see Feigenson, 2006). We have shown that people seem all too ready to accept explanations that allude to neuroscience, even if they are not accurate reflections of the scientific data, and even if they would otherwise be seen as far less satisfying. Because it is unlikely that the popularity of neuroscience findings in the public sphere will wane any time soon, we see in the current results more reasons for caution when applying neuroscientific findings to social issues. Even if expert practitioners can easily distinguish good neuroscience explanations from bad, they must not assume that those outside the discipline will be as discriminating.

Acknowledgments

We thank Paul Bloom, Martha Farah, Michael Weisberg, two anonymous reviewers, and all the members of the Cognition and Development Lab for their conversations about this work. Special thanks is also due to Marvin Chun, Marcia Johnson, Christy Marshuetz, Carol Raye, and all the members of their labs for their assistance with our neuroscience items. We acknowledge support from NIH Grant R-37-HD023922 to F. C. K.

REFERENCES

- Bloom, P. Descartes' baby. New York: Basic Books; 2004a.
Bloom P. The duel between body and soul. The New York Times 2004b:A25.

- Farah MJ. Neuroethics: The practical and the philosophical. *Trends in Cognitive Sciences* 2005;9:34–40. [PubMed: 15639439]
- Feigenson N. Brain imaging and courtroom evidence: On the admissibility and persuasiveness of fMRI. *International Journal of Law in Context* 2006;2:233–255.
- Garner R, Alexander PA, Gillingham MG, Kulikowich JM, Brown R. Interest and learning from text. *American Educational Research Journal* 1991;28:643–659.
- Garner R, Gillingham MG, White CS. Effects of “seductive details” on macroprocessing and microprocessing in adults and children. *Cognition and Instruction* 1989;6:41–57.
- Harp SF, Mayer RE. How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology* 1998;90:414–434.
- Hayes BK, Huleatt LA, Keil F. Mechanisms underlying the illusion of explanatory depth. (in preparation)
- Henson R. What can functional neuroimaging tell the experimental psychologist? *Quarterly Journal of Experimental Psychology* 2005;58A:193–233. [PubMed: 15903115]
- Keil FC. Explanation and understanding. *Annual Review of Psychology* 2006;51:227–254.
- Kelemen D. Function, goals, and intention: Children’s teleological reasoning about objects. *Trends in Cognitive Sciences* 1999;3:461–468. [PubMed: 10562725]
- Kikas E. University students’ conceptions of different physical phenomena. *Journal of Adult Development* 2003;10:139–150.
- Langer E, Blank A, Chanowitz B. The mindlessness of ostensibly thoughtful action: The role of “placebic” information in interpersonal interaction. *Journal of Personality and Social Psychology* 1978;36:635–642.
- Lombrozo T. The structure and function of explanations. *Trends in Cognitive Sciences* 2006;10:464–470. [PubMed: 16942895]
- Lombrozo T, Carey S. Functional explanation and the function of explanation. *Cognition* 2006;99:167–204. [PubMed: 15939416]
- McCabe DP, Castel AD. Seeing is believing: The effect of brain images as judgments of scientific reasoning. *Cognition*. (in press)
- Rips LJ. Circular reasoning. *Cognitive Science* 2002;26:767–795.
- Rosen J. The brain on the stand. *The New York Times Magazine* 2007 March 11;:49.
- Rozenblit L, Keil F. The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science* 2002;92:1–42.
- Shafir EB, Smith EE, Osherson DN. Typicality and reasoning fallacies. *Memory & Cognition* 1990;18:229–239.
- Trout JD. Scientific explanation and the sense of understanding. *Philosophy of Science* 2002;69:212–233.
- Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. *Science* 1974;185:1124–1131. [PubMed: 17835457]
- Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science* 1981;211:453–458. [PubMed: 7455683]

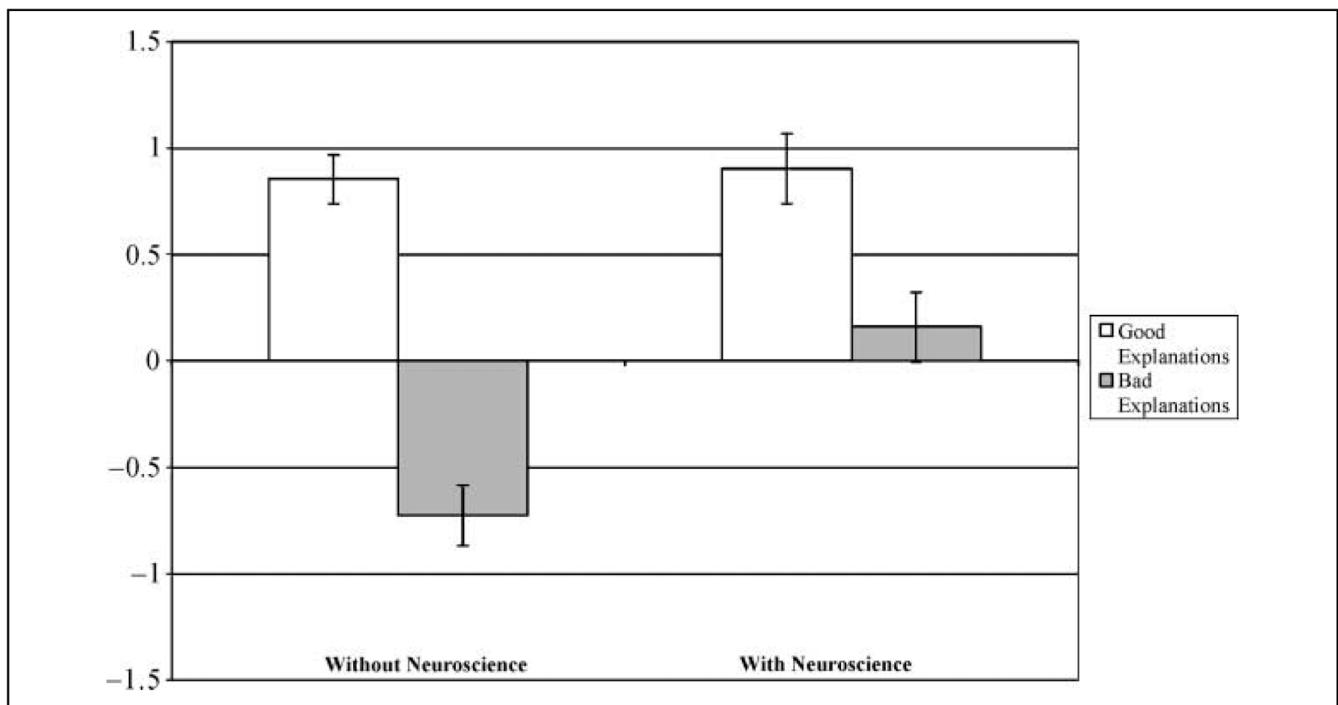


Figure 1.

Novice group. Mean ratings of how satisfying subjects found the explanations. Error bars indicate standard error of the mean.

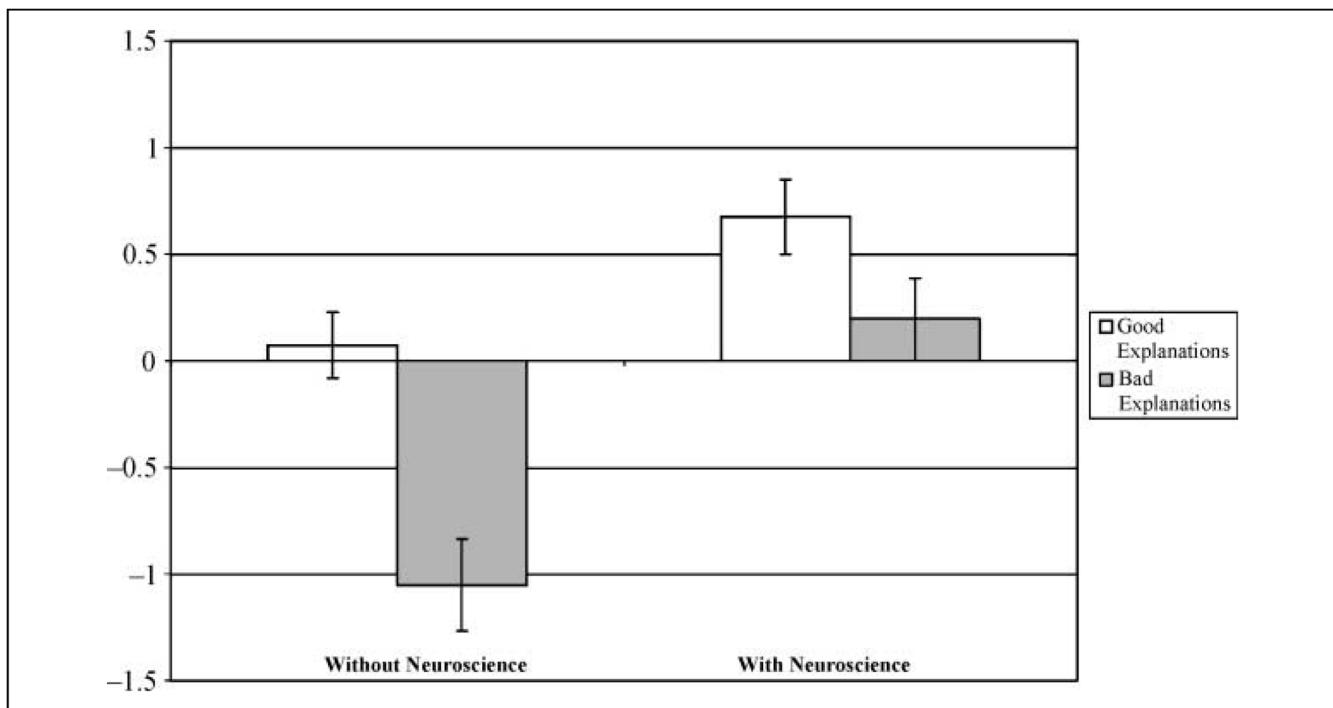


Figure 2.

Student group. Mean ratings of how satisfying subjects found the explanations. Error bars indicate standard error of the mean.

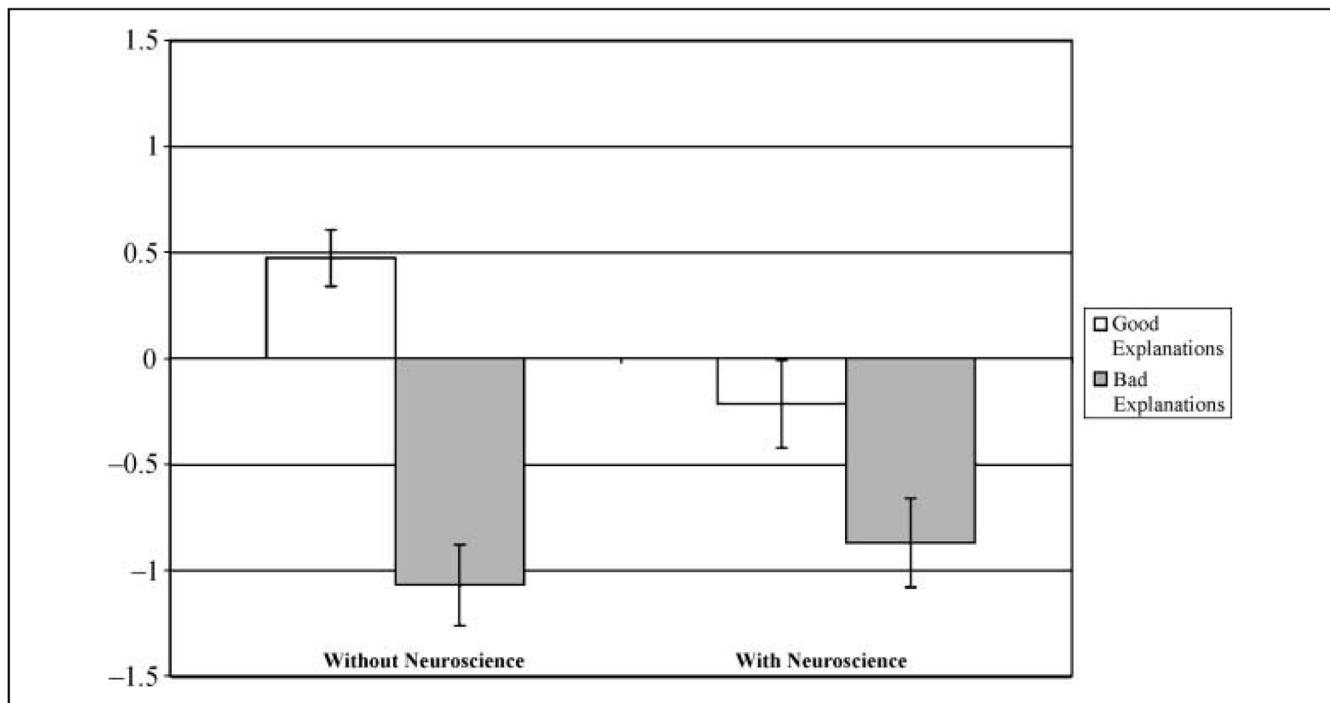


Figure 3.

Expert group. Mean ratings of how satisfying subjects found the explanations. Error bars indicate standard error of the mean.

Table 1

Sample Item

	<i>Good Explanation</i>	<i>Bad Explanation</i>
Without Neuroscience	The researchers claim that this “curse” happens because subjects have trouble switching their point of view to consider what someone else might know, mistakenly projecting their own knowledge onto others.	The researchers claim that this “curse” happens because subjects make more mistakes when they have to judge the knowledge of others. People are much better at judging what they themselves know.
With Neuroscience	Brain scans indicate that this “curse” happens because of the frontal lobe brain circuitry known to be involved in self-knowledge. Subjects have trouble switching their point of view to consider what someone else might know, mistakenly projecting their own knowledge onto others.	Brain scans indicate that this “curse” happens because of the frontal lobe brain circuitry known to be involved in self-knowledge. Subjects make more mistakes when they have to judge the knowledge of others. People are much better at judging what they themselves know.

Researchers created a list of facts that about 50% of people knew. Subjects in this experiment read the list of facts and had to say which ones they knew. They then had to judge what percentage of other people would know those facts. Researchers found that the subjects responded differently about other people’s knowledge of a fact when the subjects themselves knew that fact. If the subjects did know a fact, they said that an inaccurately large percentage of others would know it, too. For example, if a subject already knew that Hartford was the capital of Connecticut, that subject might say that 80% of people would know this, even though the correct answer is 50%. The researchers call this finding “the curse of knowledge.”

The neuroscience information is highlighted here, but subjects did not see such marking.

R. MARK WILSON, JEANNIE GAINES, AND RONALD PAUL HILL

Neuromarketing and Consumer Free Will

This article examines the impact of discoveries and methods of neuroscience on marketing practices as they relate to the exercise of individual free will. Thus, our focus centers on ethical questions involving consumers' awareness, consent, and understanding to what may be viewed as invasion of their privacy rights. After a brief introduction, the article turns to scientific literature on the brain, followed by discussion of marketing persuasion models. Ethical dilemmas within the free will paradigm and Rawlsian justice developed in moral philosophy are delineated next. The article closes with policy implications and a revised consideration of consumer privacy.

Marketers seek to influence the intricate processes of evaluation and selection by consumers, sometimes reverting to tactics and technologies that redirect decision makers without their explicit permission. Examples include product placements in videogames, movies, and television programs (see LeGresley, Muggli, and Hurt 2006). Others make use of interpersonal influences in the marketplace (McGrath and Otnes 1995; Pechmann et al. 2005). For example, marketing professionals may pay females to order specific liquors in bars or have neighbors praise particular brands of condiments or sneakers at parties (Heilbrunn 2005).

Relevant issues for our discussion are whether and to what extent marketers are willing to engage in activities that lack transparency. Few academic studies have tackled this difficult subject, providing only anecdotal evidence that the practice is more widespread than one might suspect. To address this deficit, Zinkhan, Bisessi, and Saxton (1989) asked a sample of MBA students about their willingness to deceive in a number of marketing contexts and found a broad readiness to do so in order to ensure cooperation by consumers. While the generalizability of their findings is limited, such behaviors suggest that some marketers seek to limit our understanding of their true intentions (Jeurissen and van de Ven 2006).

R. Mark Wilson is dean of the College of Business Administration at Niagara University, Niagara, NY (mwilson@niagara.edu). Jeannie Gaines is an associate professor of management at the University of South Florida St Petersburg, St Petersburg, FL (drjeanniegaines@yahoo.com). Ronald Paul Hill is the Robert J. and Barbara Naclerio Chairholder in Business and senior associate dean of Intellectual Strategy at the Villanova School of Business, Philadelphia, PA (ronald.hill@villanova.edu).

For better or for worse, opportunities to influence consumers without their full awareness may increase significantly as a result of research on brain activity. Almost twenty years ago, consumer scholars recommended using brain wave measures to study the impact of promotions on buyer behavior (see Young 2002). This perspective was controversial, especially given limitations and difficulties interpreting data from electroencephalograms (Stewart 1984, 1985). However, over this period, the disciplines of neuroscience and cognitive psychology advanced and joined forces to provide an entirely new paradigm for understanding ways consumers develop, store, retrieve, and use information (Gordon 2002). Neuroscience methodologies, especially noninvasive neuroimaging technology, now enable researchers to probe brain activity at the basic neural level of functioning (Shiv et al. 2005).

The use of data obtained from brain imaging poses ethical dilemmas for marketers. Potential moral issues emerging from neuroscience applications include awareness, consent, and understanding of individual consumers. The next section explores scientific literature on the brain, followed by a discussion of neuromarketing within models of marketing persuasion. The article then describes ethical dilemmas involving the free will paradigm argued historically in moral philosophy along with Rawlsian justice. Anticipating our results, we find that the new technology may spawn difficult ethical situations, and we offer policy implications for the future, with the intent of incorporating advantages of neuroscience within the boundaries of ethical marketing.

NEUROMARKETING AND NEUROIMAGING

The term “neuromarketing” (NM) is a recently invented moniker. *The Economist* (2004) credits Jerry Zaltman with initially proposing a union of brain-imaging technology with marketing in the late 1990s, and when the Atlanta marketing firm, BrightHouse, opened a neuromarketing division in 2001, the synthesis of neuroscience and marketing began to attract attention in science, business, and journalism. Neuromarketing has been described as “applying the methods of the neurology lab to the questions of the advertising world” (Thompson 2003, 53). Recently, the *International Journal of Psychophysiology* called neuromarketing “the application of neuroscientific methods to analyze and understand human behavior in relation to markets and marketing exchanges” (Lee, Broderick, and Chamberlain 2007, 200). Indeed, improvements in neuroimaging technologies have and will continue to advance our knowledge of how people make decisions and how marketers can influence those decisions.

The use of one noninvasive neuroimaging technology, functional magnetic resonance imaging (fMRI), has experienced especially rapid growth. fMRI enables researchers to isolate systems of neurons associated with functions of the brain. For example, when a person looks at a print advertisement, light activates some of the 125 million visual neural receptors, rods and cones, in each eye. Nerve signals travel to the midbrain, which focuses the pupils and coordinates eye movement over the advertisement. Other signals from the rods and cones pass through the optic nerve fibers, some of which cross-over to the other side of the brain so that the left half of the advertisement is perceived in the right hemisphere of the brain and the right half in the left hemisphere (Carey 2005; Dubuc 2007).

The information is processed for shape, color, and spatial location as the signals pass through the lateral geniculate nuclei on their way to assembly in the visual cortices located at the back of the brain. Memories triggered by an advertisement are stored throughout the cerebral cortex and recalled through the hippocampus located deep in each brain hemisphere; the stored emotional memories and valences are processed by the amygdala, another nerve bundle located near the base of each hemisphere (Carey 2005 Dalgleish 2004; Davidson 2003; Dubuc 2007; Kandel, Schwartz, and Jessell 2000). Using fMRI, researchers are able to image the neural activity associated with vision as well as with the cognitive and affective responses to print advertisements.

Isolating neural systems formed by the one hundred billion neurons in the human brain is a complex task. fMRI is able to locate active systems by comparing images taken of a brain performing a specific function to those of the brain when not performing that function. In an active neural system, signals travel from one neuron to another by transmitting chemical compounds, called neurotransmitters, across synapses to receptors on the receiving cell. Neurotransmitters attaching to the receptors can either facilitate or inhibit a process that will result in the firing of electrical impulses that stimulates release of neurotransmitters into synapses to the receptors of the next cell (Carey 2005; Kandel, Schwartz, and Jessell 2000). Synaptic activity of the activated network of neurons causes blood to flow to the region (Logothetis 2003; Raichle and Mintun 2006). The additional blood brings more oxygen and hydrogen to the area than is needed to replenish the system of neurons, which increases the magnetic field during a scan by a small but detectable amount (Gore 2003; Matthews and Jezzard 2004).

Improvements in hardware and software technologies continue to increase the spatial and temporal resolutions of the images, that is, the clarity of each image and the accuracy of tracking changes in brain activity over time based on these small changes in magnetic field. Current magnetic

resonance imaging machines generate a 1.5-T strong magnetic force (30,000 times the force of gravity). The protons in the nuclei of hydrogen atoms in the brain, primarily located in the blood, align their axes with this strong magnetic force. A radio wave pulse of appropriate frequency is applied at an angle to the aligned axes causing the oscillating protons to absorb energy and tip their axes away from alignment with the strong force. When the pulse ends, the particles release the absorbed energy as they return to alignment with the magnetic force. This released energy is the measured magnetic resonance signal. The information in these signals is then converted via computer software into an image of a slice of the brain. The resulting image is different from a photograph or an X-ray; it is a representation of contrasts among different tissues based on the density of the hydrogen protons and the nature of the tissue containing the protons (Detre and Wang 2002; Gore 2003; Heuttel, Song, and McCarthy 2004; Kandel, Schwartz, and Jessell 2000; Patz 2007).

During an fMRI experiment, researchers scan the individual's brain while it is not performing the function of interest, referred to as a resting brain (Raichle and Mintun 2006). Then, they perform an experiment designed to activate specific brain functions of interest while researchers quickly scan, often repeatedly, to capture changes in the signal during activity. Researchers adjust data for a myriad of factors, including the time delay between the neuronal activity and the arrival of the blood supply to the area, head movements, heartbeats, and breathing. Like a fingerprint, each brain is unique, so in studies involving more than one person, researchers "warp" each participant's brain images onto a template so that brain locations can be compared across individuals. A software program tests whether specific localities in the brain are activated during the experiment. The program colors the image of a resting brain in the locations of significant increases in blood flow, highlighting relevant networks of neurons (Brown and Semelka 1999; Gore 2003; Heuttel and McCarthy 2000; Heuttel, Song, and McCarthy 2004).

Neuroimaging and Persuasion

Researchers have applied fMRI techniques and technology to investigate the nature of decision making and persuasion. For example, Knutson et al. (2005) found the neural activity associated with calculation of expected value. They measured the brain activities of participants who were provided an informational cue about the probability and magnitude of gain or loss at the beginning of an experiment. The task was to push a button within a time limit that varied with the probability of receiving the reward. After learning the cues and the rules of this reward system, subjects entered the MRI

machine and performed 288 trials. The authors found that activation of the subcortical nucleus accumbens in the forebrain is related to magnitude of payoff but not probability of gain, while activation of the mesial prefrontal cortex is correlated with magnitude and probability of gain. These findings demonstrate that such evaluations involve both affective and cognitive neural systems.

The neuroscience literature on expected value is expanding (Breiter et al. 2001; Elliot et al. 2003), as is the larger neuroscience literature on decision making (Braeutigam 2005; Glimcher 2003; Knutson et al. 2007; Sanfey et al. 2006; Shiv et al. 2005; Zak 2004). Camerer, Lowenstein, and Prelec (2005) describe the roles of affective and cognitive processes, acting either together or separately, during decision making. The mind tags almost every concept and object with a valence that is automatically brought to mind when provoked by an appropriate symbol. Even if consumers are made aware of the affective response, it is very difficult for them to override the affective influence with cognitive reasoning. The authors speculate that cognitive processes may not be able to finalize a decision without a “go/no go” message from an affective function of the brain. Conclusions of these studies about the importance of affect in decision making parallel those of psychology and marketing (Johar, Maheswaran, and Peracchio 2006; Zajonc 1998).

Three recent articles (Braeutigam 2005; Fugate 2007; Lee, Broderick, and Chamberlain 2007) and a review of the neuroscience/marketing literature suggest that synergy between these two disciplines produced new insights into the impact of affect or emotion on the memory of visual stimuli (Ambler and Burne 1999; Ambler, Ioannides, and Rose 2000; Erk, Martin, and Walter 2005; Erk et al. 2003); antecedents of trust behavior (Ioannides et al. 2000; Fehr, Fischbacher, and Kosfeld 2005; Kosfeld et al. 2005; Zak et al. 2005); factors influencing brand selection and brand equity (Ambler et al. 2004; Braeutigam et al. 2001, 2004; du Plessis 2005; Plassmann et al. 2007); viewing time for images to enter memory (Rossiter et al. 2001; Silberstein et al. 2000); reward centers in the brain (Berns et al. 2001; Erk et al. 2002; Senior 2003); differences between evaluation of personalities and products (Yoon et al. 2006); and “branding moments” in advertisements (Young 2002). A highly publicized Coke/Pepsi fMRI study by researchers at Emory University found a significant effect of brand knowledge on brain response and expressed preference (McClure et al. 2004).

In addition to scholarly research, a number of university neuroscience programs, including those at Emory, Cal Tech, and UCLA are teaming up with private consulting firms to do applied research for large organizations such as Viacom, Kimberly-Clark, and Daimler-Chrysler (Tiltman 2005). More than 90 private neuromarketing consulting firms currently operate

in the United States as well as in an increasing number of other countries (Reid 2006). The media has sensationalized many of these investigations, alleging that marketers found the “buy button in your brain” (Dias 2006) and that the population is about to be “brain scammed” (Brain Scam? 2004). As a result, use of neuroscience in marketing has both advocates and critics. Advocates (Erk, Martin, and Walter 2005; Singer 2004; Thompson 2005) propose that the combination will allow consumers and marketers to better understand what products are desired—a win/win for both parties. Critics (Herman 2005; Huang 1998; Lovel 2003; Thompson 2003) warn that consumers’ ability to make logical, informed decisions about purchases will be compromised. Whether an advocate or a critic, many believe that neuroimaging methods will bring significant changes to marketing persuasion. Just as forty years ago when a single computer filled an entire room and its users hoped the reader would not chew up the punch cards, today’s MRI machines are large, expensive, and noisy, but it is easy to envision them, and other neuroimaging technologies developing rapidly into powerful, portable machines.

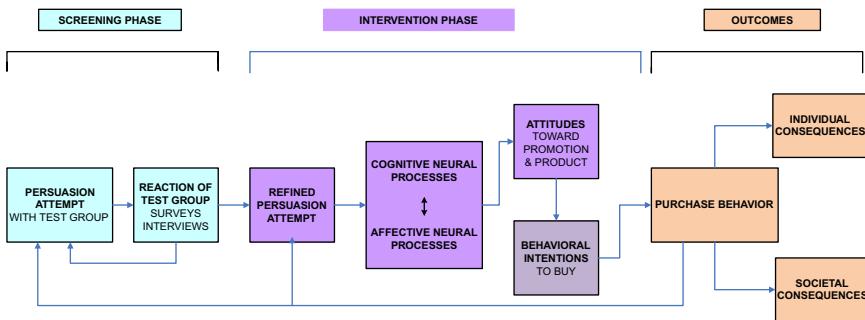
Similarly, at this point in time, conclusions drawn from the correlations between brain functions and blood flow should be viewed with caution. Their interpretation requires connecting a cognitive or affective response to neural activity, and then neural activity to a significant blood response to a region of the brain. Although neuroscientists have made significant advances in connecting neural activity to blood response, much remains to be learned about the relationship between a task-related thought or emotion and neuronal activity (Heuttel, Song and McCarthy 2004, Raichle and Mintun 2006). Nevertheless, it seems likely the new technologies will enable neuroscience and marketing researchers to better understand the role of emotions in decision making, to develop more effective methods of triggering those emotions, to build greater trust and brand loyalty, to measure intensity of an individual’s likes and dislikes, and, in general, to be more persuasive marketers. The models of marketing persuasion in the next section provide a framework for thinking about the changes that advanced neuroimaging technologies may bring.

CONSUMER PERSUASION MODELS

Traditional Consumer Persuasion Model

The Traditional Consumer Persuasion Model (Figure 1) exemplifies the way marketers have typically created more effective promotions. During the *screening* phase, a group of relevant individuals is presented with a

FIGURE 1
Traditional Consumer Persuasion Model



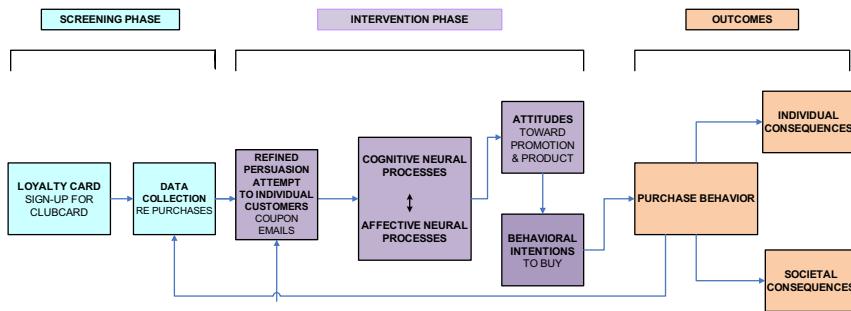
marketing stimulus, and feedback on its effectiveness is collected so that a general persuasion attempt can be refined. *Intervention* occurs when potential consumers are targeted with the resulting promotion, their cognitive and affective processes are activated, and attitudes and behavioral intentions are formed. The *outcome* phase is when a purchase occurs (or does not occur) and includes the ensuing consequences to both the individual and the larger society that can be described as positive or negative.

For example, the consumer may experience satisfaction, or even delight, with the purchase or may regret the purchase and the possible financial burden. Potential societal impacts include a boost to the overall economy or a drain if the consumer cannot pay for the item in a timely fashion. If a purchase is not made, the individual may experience regret or relief and the economy is perhaps impacted negatively in a marginal way. This model is generally accepted as an appropriate method to pursue customers and increase sales. The screening group is aware and has consented to providing feedback on various marketing stimuli. Additionally, most consumers and societal members accept this process as standard practice that does not seek to invade the private thinking and feeling of targeted consumers.

Revealed Preferences Consumer Persuasion Model

A more invasive strategy currently used by many retailers provides a second model of persuasion development based on consumer data collected at the individual level (Figure 2). For example, Tesco, Britain's largest retailer and private employer, uses a loyalty card program to record the purchasing behavior in Tesco stores of approximately 12 million UK customers and many more worldwide through Internet sales. Consumers willingly disclose personal information required for the Clubcard because

FIGURE 2
Revealed Preferences Consumer Persuasion Model



“points” based on the cardholder’s total purchases can be redeemed for discounts on future purchases or for air miles in frequent-flier programs. Dunnhumby, a marketing research firm owned primarily by Tesco, analyzes customer data by correlating characteristics of products that an individual buys with those of other people with similar, but not identical, purchases and shopping habits. Tesco marketers are then able to design promotions that cater to specific clusters of individuals such as targeted e-mails and quarterly mailings. These communications include coupons for items the individual typically buys as well as for items she/he is likely to buy based on data analyses.

Tesco and others have been remarkably successful using customer information to increase their own sales and by selling information about purchasing behavior from the dataset to other retailers (Humby, Hunt, and Phillips 2007; Rigby 2006; Rohwedder 2006). In the Revealed Preferences Consumer Persuasion Model, the *screening* phase consists of signing customers up for the Clubcard as well as collecting and analyzing data on each purchase. In the *intervention* phase, the refined persuasion attempt is crafted through specific e-mails, coupons based on consumer preferences, and in the future via grocery carts with small LCD screens containing advertisements targeted to the individual shopper. Cognitive and affective responses, attitude formation and behavioral intentions, and purchase decisions unfold consistent with *outcomes* of the former model.

Information about purchasing decisions is fed back to marketers to refine succeeding persuasion attempts. As with the Traditional Model, engaged individuals are aware and consent to what might be deemed under other circumstances as an invasion of their privacy, at least during the screening phase. However, it is doubtful that the majority understands the extent of the statistical manipulations of their personal data that inform the

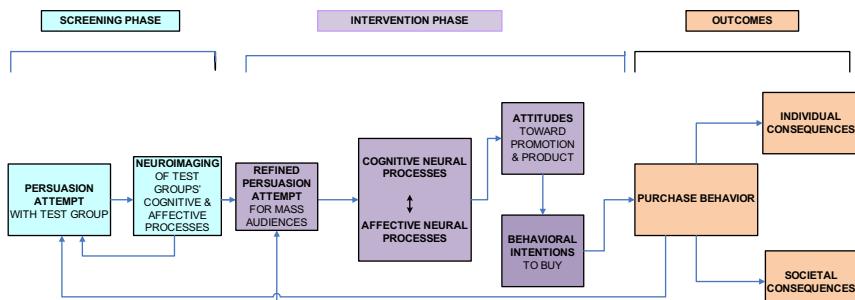
intervention phase, even though it is recounted in a recent book titled *Scoring Points* (Humby, Hunt, and Phillips 2007). The major difference between the Revealed Preferences Model and the Traditional Model is that customer preference data are collected and used to target the individual as well as others “like” him or her. The feedback obtained through purchases is also much more specific and accurate. Tesco’s stated marketing goal is to give people what they want, and they have been creative in the use of data to determine preferences.

Collective Neuromarketing Persuasion Model

The Collective Neuromarketing Persuasion Model (Figure 3) differs from the first two models only in the *screening* phase. As opposed to study groups or loyalty cards, this model introduces neuroimaging techniques into the consumer behavior paradigm. Here, a subset of consumers agrees to neuroimaging measurement while observing various marketing stimuli. The new measurement methods record important nonconscious affective influences, and the results are then used to design future persuasion attempts.

During the *intervention* phase, the refined persuasion attempt is based on brain scan data of the test group and is presented to future potential buyers in relevant settings. The stimulus is processed by consumers, through cognitive and affective mechanisms, to form an attitude toward the brand or product. If neuromarketers are successful triggering affective areas of the brain associated with rewards or pleasure, the consumer develops a positive attitude toward the product, forms a behavioral intention to buy, and ultimately purchases the item in question (*outcome* phase). The purchase behavior forms a feedback loop to the screening phase where persuasion

FIGURE 3
Collective Neuromarketing Consumer Persuasion Model



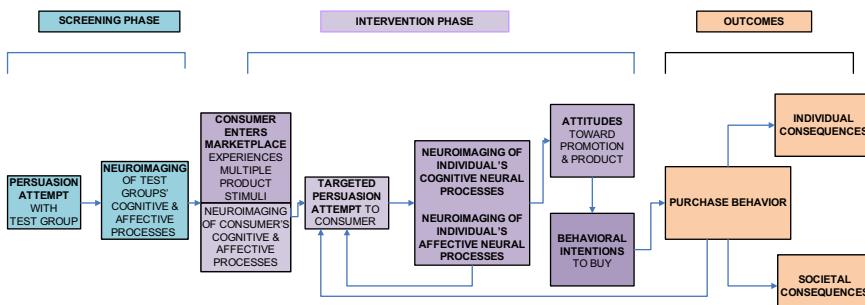
attempts are continuously refined. This process allows for constant revision of marketing stimuli based on a combination of brain imaging in tandem with actual consumer behavior. As in the two previous models, the screening phase is conducted with participants' awareness and consent, with the major difference being the invasiveness of data collection on consumer reactions to promotions through brain scans and their subsequent usages.

Individual Neuromarketing Persuasion Model

The Individual Neuromarketing Model (Figure 4) is a look at the possibilities that may exist in the coming years. As with the collective neuromarketing model, the *screening* phase consists of neuroimaging used with a test group of consumers. However, the *intervention* phase in this model is directed only at individuals as opposed to an undifferentiated mass of consumers. For instance, consider a buyer who enters a marketplace such as a department store or mall where she/he typically is bombarded with marketing stimuli. In order to better understand its impact, retailers may neuro-screen potential customers upon entering, registering reactions to what they see, hear, feel, touch, taste, and/or smell, and combining these measurements and outcomes with previous readings based on earlier visits.

As a consequence, marketing attempts could be targeted directly to consumers based upon their brain scans. For example, if neuroimaging data suggest a positive response to the touching of jewelry, the consumer may experience a personalized discount prominently displayed in their sightline in order to provide encouragement for purchase. While subjected to these specific persuasion attempts, the individual's brain is continuously monitored to determine if the stimuli are having the desired effects. When the transaction is or is not completed, the results might be fed back and

FIGURE 4
Individual Neuromarketing Consumer Persuasion Model



recorded to create an increasingly more sophisticated picture of the consumer.

The Individual Neuromarketing Model suggests the greatest concerns about personal awareness and consent, barring some form of voluntary or government-mandated disclosures like the ones occurring now on genetically modified foods. Additionally, while marketer use of neuroimaging technology will allow consumers to experience exceptionally accurate and effective marketing stimuli, concerns exist about how individuals' privacy will be maintained, who ultimately owns brain scans, whether scans can be sold to other persons or institutions, and what happens to extraneous information, such as health problems, revealed by the scans. Such issues are indicative of both possibilities and dilemmas that lie ahead at the intersection of marketing and neuroscience.

Together, persuasion models suggest new forms of consumer miscomprehension that may lead to additional privacy concerns. While a burgeoning literature is developing on topics such as neuroethics, its primary focus is on applications outside the marketing domain. The next section attempts to fill this void by bringing a unique perspective to neuroethics within a consumer-driven context that is organized around the concept of free will. This philosophical premise is described briefly, and the resulting argument frames the ethical implications caused by neuroscience thinking and practice. Models are evaluated by uniform criteria, followed by closing remarks that signal the broader policy implications that may become important as our understanding and training advances.

NEUROSCIENCE, FREE WILL, AND PERSUASION ATTEMPTS

Earlier discussion of neuroscience shows that our biology has an overwhelming impact on decision making and action, suggesting that even morality may be outside our purview (see Fukuyama 2002 for an excellent discussion). Implicit to this belief is that *knowing* what portions of our brains are stimulated may *reveal* the nature of resulting behaviors. Also noted previously, technology necessary to create visual and dynamic representations of such processes is developing rapidly, and machinery that is both portable and unobtrusive may soon be available for use by researchers and marketers. Such equipment could allow monitoring of consumers with or without their awareness, permission, or understanding.

This "brave new world" (Huxley 1932) begs the question as to appropriate responsibilities between consumers and those parties that seek to influence their beliefs, feelings, and behaviors. Among philosophers there are differences of opinion on the nature and primacy of human beings that

may inform this debate (see Klemke 2000). At one end of the spectrum are scholars who believe that all living creatures are similar, with some having a few distinctive features but still operating by instinct (Flanagan 2002). Other researchers who cross the boundaries between science and ethics recognize these genetic predilections yet believe in our ability to rise above biology. (Once again see the review by Fukuyama 2002.) Part of this distinction is based on the long-standing debate concerning free will and its role in our transcendence beyond nature (Baggini 2005).

The concept of the freedom of the will moved front and center during the Age of Enlightenment in eighteenth-century Europe, when philosophers argued about our capacity to use rational judgments to determine both truth and moral behavior (Wallerstein 1997). Succeeding generations of scholars examined various aspects of this construct, often suggesting that our culpability in situations is dependent upon making conscious choices among the variety of options available and acting voluntarily (Hospers 1953; Spence 1996). A term coined to represent this context is *uncaused causer* (Greene and Cohen 2004), which recognizes that current behavior is not perceived to be controlled by anyone or anything external to the decision maker (Levy 2003). As a consequence, free will provides a basis upon which people have sought to differentiate themselves from each other and exert that their lives have real importance. Modern approaches eschew beliefs in externally imposed meaning in favor of internally generated yearning (Baggini 2005).

Applied ethics scholars have used such theories in business/marketing contexts to provide a normative structure whereby actions and outcomes in exchange relationships can be judged (see Murphy, Lacznak, and Wood 2007). A complementary approach to free will is contractualism, and the work of Rawls (1971) is at the centerpiece of its applications (see Brock 1998; Toenjes 2002). His frame provides legitimate standards by which the distribution of rights and responsibilities can be determined to the consensual agreement of exchange partners. These agreements are based on individual dignity that social arrangements should not violate.

Inherent to this perspective is the division between inequities that result from poor decision making and inequalities that are due to conditions beyond one's control (Tan 2001). For example, differences in relative power, resources, or information based on dissimilarities in effort or contribution to exchange relationships are morally acceptable (Cohen 1997). However, inequities due to discrimination, selfishness, or other forms of unjustifiable external constraints clearly are immoral. As a whole, Rawlsian justice suggests that social actors must find ways to interact that satisfy these conditions and produce solutions that are acceptable to all parties

(Zanetti 2001). In the final analysis, Rawls (1971) believes that rational people will establish systems of exchange that avoid downside risks associated with poor starting positions and allow for fair allocations.

Our contention is that neuroscience findings and methods hold the potential for marketing practices that threaten consumers' abilities to follow preferences and dictates according to free will (Greene 2003) and contradict Rawlsian justice. This context suggests that external constraints on decision making imposed by applications of neural manipulation are possible violations. Transgressions are particularly troublesome when manipulation occurs without explicit awareness, consent, and understanding. The next subsections examine the ethical issues that arise in company-to-customer communications. Potential dilemmas are delineated using the models described previously as the frame of reference, and disruption of the will advances as representations move from the traditional to neuromarketing models. Concerns related to screening, intervention, and outcome phases are presented using language involving the exercise of choice.

Ethical Issues for Traditional and Revealed Preference Models

The Traditional Model follows the more conventional path in the development and dissemination of marketing communications for mass audiences. Advertisements or other persuasion attempts are assessed using a variety of techniques, including paper and pencil or baseline physiological measures. While potential ethical conflicts may arise, the primary practice is that test consumers are aware of and consent to these assessments prior to and during exposure to marketing stimuli in the screening phase. Lack of transparency may occur, for example, in the use of one-way mirrors or other forms of unobtrusive observation of reactions, but such procedures typically involve behaviors in a more public setting and therefore may not necessarily be viewed as violating individual privacy rights.

The same perspective may be true of the Revealed Preferences Model whereby consumers willingly disclose a host of private information about themselves in what they believe to be reciprocal relationships with firms. These data often are used in subsequent persuasion attempts that are targeted directly at individual consumers. Since these persons have agreed to this arrangement by virtue of their participation, it can be assumed that they willingly acknowledge and accept use of their profiles in ways that expand opportunities for them to do business with focal retailers or other involved marketers.

Nonetheless, the free will frame presented earlier suggests possible ethical violations that are a function of the lack of true awareness and consent

on the part of subjects and targeted consumers. Even under the Traditional Model, it is unlikely that participants in development of various marketing stimuli fully understand the uses of information gleaned during the screening phase and how they might be used during future persuasion attempts. Additionally, while the responses provided are typically applied in summary form only, the rights to use this information pass to the agency or firm without an informed assessment of potential consequences by test consumers. The outcomes of data manipulation and usage are considerably greater for the Revealed Preferences Model since information is more likely to be of a sensitive nature, sold to third-party marketers, and used to profile specific consumers without even cursory awareness.

Ethical Issues for the Collective Neuromarketing Persuasion Model

The Collective Neuromarketing Model also follows the traditional path involved in the creation of marketing communications for targeted consumers. The primary difference is that neuroimaging technology is used during screening of persuasion attempts, which represents a quantum change in marketers' ability to judge the impact of communications relative to measures discussed under the previous models. Not only does neuroimaging allow researchers to "read the minds" of test subjects more accurately, it also permits them to delineate which stimuli trigger excitement, trust, pleasure, i.e., the emotions that lead people to buy. To the extent these stimuli are unrelated to product characteristics, the result is an attempt to manipulate the consumer's purchase decision.

Another area of concern is the degree to which test subjects understand fully the personal nature of brain scans that are now property of a marketing group or organization. If the research protocol leaves the test subjects unaware of potential privacy issues, such lack of transparency may jeopardize intimate neurological data.

After screening is finished, marketing managers begin a controlled release of stimuli into the marketplace designed to influence cognitive and affective neural processes of consumers. Once again, a quick inspection of these procedures suggests similarity with current marketing practices. However, the underlying intent is to trigger emotions that encourage purchase rather than to provide consumers with accurate information on which to make beneficial decisions.

The free will frame presented earlier suggests that the primary ethical violation is a function of the lack of awareness, consent, and understanding on the part of targeted consumers. Given these conditions, potential customers are unable to make informed decisions about the extent to which

they would choose to be influenced by such marketing stimuli. Some scholars may contend that this problem exists with *all* persuasion attempts since they often are placed in our sensory path without tacit permission and the strategic intentions of their developers remain unknown. Nonetheless, a fundamental distinction between other marketing and collective neuromarketing tactics is that the former attempts to change beliefs, attitudes, and behaviors through well-recognized means, while the latter are expert attempts to trigger buying emotions in consumers.

Ethical Issues for the Individual Neuromarketing Persuasion Model

The Individual Neuromarketing Model mirrors the possible ethical issues associated with the screening phase of the previous models; however, the similarities end there. Once the range of possible neural reactions are explored fully, potential customers are exposed to marketing stimuli with the intent of creating an individual profile for the purpose of manipulation using a running series of fine-tuned persuasion attempts that are continuously monitored and recorded. The first ethical dilemma that arises within the intervention phase concerns whether consumers are aware of and consent to omnipresent scrutiny and to targeted/personal exposure to marketing stimuli. The worst case scenario involves the use of neuroimaging technology in public contexts where consumers would be oblivious to its employ and/or its resulting effects on brain functioning and decision making. Such a context limits consumer free will and violates Rawlsian ethics since a rational person would never select to be so manipulated.

The next logical scenario allows for awareness of monitoring and development of personalized marketing tactics using neuroscientific methods and technologies without true consent. While this combination may seem unlikely, the possibility exists that consumers will agree to enter a public shopping environment where they undergo screening in order to maintain access to marketplace activities that are not easily found elsewhere. Thus, their perceived or real consumption restrictions may cause potential customers to subject themselves to unwanted invasion of private mental processes and to bombardment of their personal space with intrusive marketing stimuli. These concerns also exist with the previous scenario, but awareness may reduce the possibility of manipulation, leaving ethical violations associated with obligatory consent rather than ignorance of intent.

The final scenario applicable to this model includes situations where consumers are aware of and consent to scrutiny and persuasion attempts. This situation eliminates many of the dilemmas noted with the two previous scenarios, but a few issues remain. First, agreement does not ensure a

complete understanding of how personalized targeting will impact buyer behavior, likely necessitating warning systems and social marketing programs that are currently used for addictive or complex products such as tobacco and alcohol or financial services and healthcare. A second problem becomes one of relative quality of consumption of persons without these opportunities. Given the inequities in our material world, some consumers are likely to experience vulnerability because of their lack of access to such technologies. The third issue involves how beneficial a resulting purchase is to the consumer. When a consumer purchases a product based on a decision in which marketing stimuli unrelated to product characteristics cause affective neural systems to override cognitive processes, the final purchase outcome may not always be in the best interest of the consumer.

CLOSING REMARKS

This investigation brings disparate literature and secondary research together in order to explore the complex persuasion environment for marketers and consumers of their goods and services resulting from neuroscientific discoveries. Ethical dilemmas are exacerbated by use of neuromarketing methods and data, and center on issues of consumer free will and privacy. The ability to exercise free will in purchasing decisions is informed by Preston's (2002) discussion of problematic "antifactual" advertising content consisting of puffery, obvious false claims, and lifestyle claims. While not technically considered "deceptive advertising," by the Federal Trade Commission, they clearly fail to inform consumers about products—ostensibly the basis of rational purchasing decisions. Neurotechnology enables marketers to refine persuasion attempts using noninformative or misinformative content, with the potential to trigger very positive affective responses in consumers. While some may argue that this technique only encourages consumers to buy *what* they really want, Rotfeld (2007) questions the whole premise of selling people *only* what they want. He suggests that marketing should be "going beyond giving consumers what they like," but rather "helping more people understand what they really should want" (p. 384) or need. This stance speaks to the importance of marketplace education so consumers can exercise free will around purchasing decisions based on accurate information.

Issues of awareness, consent, and understanding form a cohesive set of moral questions that are addressed, in part, by free will and Rawlsian justice. For example, behind a "veil of ignorance" where one fails to know whether she/he is the marketer or consumer, would she/he select to be oblivious, ignorant, or restricted? The answer is a clear "no" from the

perspective of any individual looking out for her/his best interest, and for whom the ability to exercise free will is a high priority.

Unfortunately, self-regulation and public policy lag behind current practice and future opportunities. Just as copyright and varied intellectual property laws established prior to the Internet fail to serve existing legal needs, so our thinking about lack of transparency surrounding promotional activities should be updated to include neuromarketing methods. Many within the academic and practitioner communities may suggest that the natural skepticism of consumers developed over centuries of dealings in the marketplace will provide a natural barrier to potential harm. Nonetheless, distrust is only activated in ways that are relevant to accumulated experiences, and this "brave new world" portends new transparency concerns that may have insidious effects as well as unknown consequences.

Regardless, the potential restriction of free will and privacy invasiveness enabled by neuroimaging technology requires attention by governmental and academic constituencies. The rapid collection, assessment, and deployment of brain scanning data anticipated by the latter models reveal new terrain for researchers and legislators interested in the protection of consumer rights. Questions as to who owns such information, how it may be combined with other databases in order to develop more sophisticated and targeted marketing efforts, and under what conditions it may be sold or traded with others represent areas that will require attention. The Federal Trade Commission standards, as articulated in their *Fair Information Practice Principles*, are a good starting point and are designed to acknowledge the rights of consumers (www.ftc.gov/reports/privacy3/fairinfo.shtm).

These principles are built around five core ideals. The first is notice/awareness and is central to the remaining standards. Under this guiding principle, consumers should be told who is collecting data, its possible uses, and any potential recipients. The second is choice/consent, which is consistent with our previous discussion on neuromarketing. Consumers are given the opportunity to opt-in or opt-out of the collection of information and also have the ability to tailor the nature of their data and its uses. The third is access/participation, which is concerned with the consumer's capacity to view, verify, and contest the completeness and accuracy of information about them in a timely and efficient way. The fourth is integrity/security and requires that marketers, and their firms ensure that data are up to date and protected against unauthorized access or manipulation.

The fifth principle involves enforcement/redress. Given our free will premise, marketing practitioners should be expected to communicate the uses and outcomes of neuroimaging technology prior to consumer

exposure, to allow individuals to opt-out of any or all aspects of the collection process without penalty, to provide easily accessible and understandable feedback on personal information, and to ensure that appropriate safeguards are in place to prevent unwanted third-party exposure. The first line of defense is self-regulation and would require a cross-disciplinary group of scholars and practitioners to come together to develop standards, assessment mechanisms, and sanctions. If this fails to resolve the most serious problems, private remedies through the court system may establish the criteria upon which neuromarketing activities will be judged and constrained, leading to legislative solutions and lawmaker control.

In conclusion, the issues of freedom of will, privacy rights, and the development and dissemination of advertisements by business operations are broadened significantly by the inclusion of neuroscience methods and findings. The conjoining of marketing and neuroscience clearly is in its infancy, and only the Collective Neuromarketing Model is in use by a growing assortment of scholars and practitioners. Yet, adoption of the Individual Neuromarketing Model is more than musings in postmodern novels. It represents possibilities that will need a combination of voluntary compliance and regulatory oversight in order to avoid some of the dilemmas noted here. A critical role for policy makers and consumer scholars is to inform this debate by monitoring the latest neuroscientific findings and evaluating their implications for ethical marketing practice.

REFERENCES

- Ambler, Tim, Sven Braeutigam, John Stins, Steven Rose, and Stephen Swithenby. 2004. Salience and Choice: Neural Correlates of Shopping Decisions. *Psychology and Marketing*, 21 (4): 247–261.
- Ambler, Tim and Tom Burne. 1999. The Impact of Affect on Memory of Advertising. *Journal of Advertising Research*, 39 (March): 25–34.
- Ambler, Tim, Andreas Ioannides, and Steven Rose. 2000. Brands on the Brain: Neuro-Images of Advertising. *Business Strategy Review*, 11 (3): 17–30.
- Baggini, Julian. 2005. *What's It All About? Philosophy and the Meaning of Life*. New York: Oxford University Press.
- Berns, Gregory, Samuel McClure, Giuseppe Pagnoni, and Read Montague. 2001. Predictability Modulates Human Brain Response to Reward. *Journal of Neuroscience*, 21 (8): 2793–2798.
- Braeutigam, Sven. 2005. Neuroeconomics—From Neural Systems to Economic Behavior. *Brain Research Bulletin*, 67 (5): 355–360.
- Braeutigam, Sven, Steven Rose, Stephen Swithenby, and Tim Ambler. 2004. The Distributed Neuronal Systems Supporting Choice-Making in Real-Life Situations: Differences between Men and Women When Choosing Groceries Detected using Magnetoencephalography. *European Journal of Neuroscience*, 20 (1): 293–302.
- Braeutigam, Sven, J.F. Stins, Steven Rose, Stephen Swithenby, and Tim Ambler. 2001. Magnetoencephalographic Signals Identify Stages in Real-Life Decision Processes. *Neural Plasticity*, 8 (4): 241–254.
- “Brain Scam?” 2004. *Nature Neuroscience*, 7 (7): 683.

- Breiter, Hans C., Itzhak Aharon, Daniel Kahneman, Anders Dale, and Peter Shizgal. 2001. Functional Imaging of Neural Responses to Expectancy and Expense of Monetary Gains and Losses. *Neuron*, 30 (May): 619–639.
- Brock, Gillian. 1998. Are Corporations Morally Defensible? *Business Ethics Quarterly*, 8 (4): 703–721.
- Brown, Mark A., and Richard C. Semelka. 1999. *MRI Basic Principles and Applications*. New York: Wiley-Liss.
- Camerer, Colin, George Lowenstein, and Drazen Prelec. 2005. Neuroeconomics: How Neuroscience Can Inform Economics. *Journal of Economic Literature*, 43 (March): 9–64.
- Carey, Joseph, ed. 2005. *Brain Facts*. Washington, DC: Society for Neuroscience.
- Cohen, G. 1997. Where the Action Is: On the Site of Distributive Justice. *Philosophy and Public Affairs*, 26 (Winter): 3–30.
- Dalgleish, Tim. 2004. The Emotional Brain. *Nature Reviews Neuroscience*, 5 (7): 582–589.
- Davidson, Richard J. 2003. Darwin and the Neural Bases of Emotion and Affective Style. *Annals of the New York Academy of Sciences*, 1000 (December): 316–336.
- Detre, John A., and Jiongjiong Wang. 2002. Technical Aspects and Utility of fMRI Using BOLD and ASL. *Clinical Neurophysiology*, 113 (5): 621–634.
- Dias, David. 2006. A ‘Buy Button’ in Your Brain? *National Post*, July 1.
- du Plessis, Erik. 2005. *The Advertised Mind*. London, UK: Millward Brown and Kogan Page Limited.
- Dubuc, B. 2007. The Brain from Top to Bottom. Canadian Institutes of Health Research: Institute of Neurosciences, Mental Health and Addiction. http://www.thebrain.mcgill.ca/flash/index_i.html (retrieved January 30, 2007).
- Elliot, Rebecca, Jana L. Newman, Olivia A. Longe, and J.F. William Deakin. 2003. Differential Response Patterns in the Striatum and Orbitofrontal Cortex to Financial Reward in Humans: a Parametric Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, 23 (1): 303–307.
- Erk, Susanne, Markus Kiefer, J. Grothe, Arthur Wunderlich, Manfred Spitzer, and Henrik Walter. 2003. Emotional Context Modulates Subsequent Memory Effect. *NeuroImage*, 18 (2): 439–447.
- Erk, Susanne, Sonja Martin, and Henrik Walter. 2005. Emotional Context During Encoding of Neutral Items Modulates Brain Activation Not Only During Encoding but also During Recognition. *NeuroImage*, 26 (3): 829–838.
- Erk, Susanne, Manfred Spitzer, Arthur Wunderlich, Lars Galley, and Walter Henrik. 2002. Cultural Objects Modulate Reward Circuitry. *NeuroReport*, 13 (18): 2499–2503.
- Fehr, Ernst, Urs Fischbacher, and Michael Kosfeld. 2005. Neuroeconomic Foundations of Trust and Social Preferences: Initial Evidence. *American Economics Association Papers and Proceedings*, 95 (2): 346–351.
- Flanagan, Owen. 2002. *The Problem of the Soul*. New York: Basic Books.
- Fugate, Douglas. 2007. Neuromarketing: A Layman’s Look at Neuroscience and Its Potential Application to Marketing Practice. *Journal of Consumer Marketing*, 24 (7): 385–394.
- Fukuyama, Francis. 2002. *Our Posthuman Future: Consequences of the Biotechnology Revolution*. New York: Farrar, Straus, and Giroux.
- Glimcher, Paul W. 2003. *Decisions, Uncertainty, and the Brain*. Cambridge, MA: MIT Press.
- Gordon, Wendy. 2002. The Darkroom of the Mind: What Does Neuropsychology Now Tell Us About Brands? *Journal of Consumer Behaviour*, 1 (February): 280–292.
- Gore, John C. 2003. Principles and Practice of Functional MRI of the Human Brain. *Journal of Clinical Investigation*, 112 (1): 4–9.
- Greene, Joshua. 2003. From Neural ‘Is’ to Moral ‘Ought’: What are the Moral Implications of Neuroscientific Moral Psychology? *Nature Reviews Neuroscience*, 4 (October): 846–850.
- Greene, Joshua and Jonathan Cohen. 2004. For the Law, Neuroscience Changes Nothing and Everything. *Philosophical Transactions of the Royal Society of London B*, 359: 1775–1785.
- Heilbrunn, Jacob. 2005. Totally Fake Memo: Covert Marketing. *Los Angeles Times*, January 15.
- Herman, Steve. 2005. Selling to the Brain. *Global Cosmetic Industry*, 173 (5): 64–66.
- Heuttel, Scott A. and Gregory McCarthy. 2000. Evidence for a Refractory Period in the Hemodynamic Response to Visual Stimuli as Measured by MRI. *NeuroImage*, 11 (5): 547–553.

- Heuttel, Scott A., Allen W. Song, and Gregory McCarthy. 2004. *Functional Magnetic Resonance Imaging*. Sunderland, MA: Sinauer Associates, Inc.
- Hospers, John. 1953. Meaning and Free Will. *Philosophy and Phenomenological Research*, 10 (March): 307–330.
- Huang, Gregory. 1998. The Economics of Brains. *Technology Review*, 108 (5): 74–76.
- Humby, Clive, Terry Hunt, and Tim Phillips. 2007. *Scoring Points*. London, UK: Kogan Page Limited.
- Huxley, Aldous. 1932. *Brave New World*. New York: Harper/Perennial.
- Ioannides, Andreas, Lichan Lui, Dionyssios Theofilou, Jurgen Dammers, Tom Burne, Tim Ambler, and Steven Rose. 2000. Real Time Processing of Affective and Cognitive Stimuli in the Human Brain Extracted from MEG Signals. *Brain Topography*, 13 (1): 11–19.
- Jeurissen, Ronald and Bert van de Ven. 2006. Review Article: Developments in Marketing Ethics. *Business Ethics Quarterly*, 16 (3): 427–439.
- Johar, Gita Venkataramani, Durairaj Maheswaran, and Laura A. Peracchio. 2006. MAPping the Frontiers: Theoretical Advances in Consumer Research on Memory, Affect, and Persuasion. *Journal of Consumer Research*, 33 (1): 139–149.
- Kandel, Eric R., James H. Schwartz, and Thomas M. Jessell. 2000. *Principles of Neural Science*. 4th edition. New York: McGraw-Hill.
- Klemke, E.D. 2000. *The Meaning of Life*. New York: Oxford University Press.
- Knutson, Brian, Scott Rick, G. Elliot Wimmer, Drazen Prelec, and George Loewenstein. 2007. Neural Predictors of Purchases. *Neuron* 53 (January 4): 147–156.
- Knutson, Brian, Jonathan Taylor, Matthew Kaufman, Richard Peterson, and Gary Glover. 2005. Distributed Neural Representation of Expected Value. *Journal of Neuroscience*, 25 (19): 4806–4812.
- Kosfeld, Michael, Markus Heinrichs, Paul Zak, Urs Fischbacher, and Ernst Fehr. 2005. Oxytocin Increases Trust in Humans. *Nature*, 435 (June 2): 673–676.
- Lee, Nick, Amanda Broderick, and Laura Chamberlain. 2007. What is ‘Neuromarketing’? A Discussion and Agenda for Future Research. *International Journal of Psychophysiology*, 63 (2): 199–204.
- LeGresley, Eric M., Monique E. Muggli, and Richard D. Hurt. 2006. Movie Moguls: British American Tobacco’s Covert Strategy to Promote Cigarettes in Eastern Europe. *European Journal of Public Health*, 16 (5): 505–508.
- Levy, Daniel A. 2003. Neural Holism and Free Will. *Philosophical Psychology*, 16 (2): 205–226.
- Logothetis, Nikos K. 2003. The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal. *Journal of Neuroscience*, 23 (10): 3963–3971.
- Lovel, Jim. 2003. Nader Group Slams Emory for Brain Research. *Atlanta Business Chronicle*, December 8.
- Matthews, P.M. and P. Jezzard. 2004. Functional Magnetic Resonance Imaging. *Journal of Neurology, Neurosurgery, and Psychiatry*, 75 (1): 6–12.
- McClure, Samuel M., Jian Li, Damon Tomlin, Kim S. Cypert, Latane M. Montague, and P. Read Montague. 2004. Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron*, 44 (October): 379–387.
- McGrath, Mary Ann and Cele Otnes. 1995. Unacquainted Influencers: When Strangers Interact in the Retail Setting. *Journal of Business Research*, 32 (3): 261–272.
- Murphy, Patrick, Gene Lacziak, and Graham Wood. 2007. An Ethical Basis for Relationship Marketing: A Virtue Ethics Perspective. *European Journal of Marketing*, 41 (1/2): 37–57.
- Patz, Sam. 2007. MR Signal Sources. In *The Whole Brain*, edited by Keith A. Johnson. <http://www.med.harvard.edu/AANLIB/sigors.html> (retrieved January 30, 2007).
- Pechmann, Cornelia, Linda Levine, Sandra Loughlin, and Frances Leslie. 2005. Impulsive and Self-Conscious: Adolescents’ Vulnerability to Advertising and Promotion. *Journal of Public Policy & Marketing*, 24 (Fall): 202–221.
- Plassmann, Hilke, Tim Ambler, Sven Braeutigam, and Peter Kenning. 2007. What Can Advertisers Learn from Neuroscience? *International Journal of Advertising*, 26 (2): 151–175.
- Preston, Ivan. 2002. A Problem Ignored: Dilution and Negation of Consumer Information by Antifactual Content. *The Journal of Consumer Affairs*, 36 (2): 263–283.

- Raichle, Marcus E. and Mark A. Mintun. 2006. Brain Work and Brain Imaging. *Annual Review of Neuroscience*, 29 (July): 449–476.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Reid, Alasdair. 2006. MRI Scanners Can Improve Advertising Effectiveness. *The Economic Times*, January 19.
- Rigby, Elizabeth. 2006. Eyes in the Till. *Financial Times*, November 11.
- Rohwedder, Cecilia. 2006. Stores of Knowledge. *Wall Street Journal*, Eastern Edition, June 6.
- Rossiter, John, Richard Silberstein, Geoff Nield, and Philip Harris. 2001. Brain-Imaging Detection of Visual Scene Encoding in Long-term Memory for TV Commercials. *Journal of Advertising Research*, 41 (2): 13–21.
- Rotfeld, Herbert Jack. 2007. Mistaking a Marketing Perspective for Ethical Analysis: When Consumers Can't Know That They Should Want. *Journal of Consumer Marketing*, 24 (7): 383–384.
- Sanfey, Alan G., George Loewenstein, Samuel M. McClure, and Jonathan D. Cohen. 2006. Neuroeconomics: Cross-currents in Research on Decision-making. *Trends in Cognitive Sciences*, 10 (3): 108–116.
- Senior, Carl. 2003. Beauty in the Brain of the Beholder. *Neuron*, 38 (4): 525–528.
- Shiv, Baba, Antoine Bechara, Irwin Levin, Joseph W. Alba, James R. Bettman, Laurette Dube, Alice Isen, Barbara Mellers, Ale Smidts, Susan J. Grant, and A. Peter McGraw. 2005. Decision Neuroscience. *Marketing Letters*, 16 (3/4): 375–386.
- Silberstein, R.B., P.G. Harris, G.A. Nield, and A. Pipingas. 2000. Frontal Steady-State Potential Changes Predict Long-term Recognition Memory Performance. *International Journal of Psychophysiology*, 39 (1): 79–85.
- Singer, Emily. 2004. They Know What You Want. *New Scientist*, July 31, <http://www.newscientist.com>.
- Spence, Sean A. 1996. Free Will in the Light of Neuropsychiatry. *Philosophy, Psychiatry, & Psychology*, 3 (2): 75–90.
- Stewart, David. 1984. Physiological Measurement of Advertising Effects. *Psychology & Marketing*, 1 (1): 43–48.
- Stewart, David. 1985. Differences between Basic Research and the Validation of Specific Measures: A Reply to Weinstein et al. *Psychology & Marketing*, 2 (1): 41–49.
- Tan, Kok-Chor. 2001. Critical Notice of John Rawls' *The Law of Peoples*: With the 'Idea of Public Reason' Revisited. *Canadian Journal of Philosophy*, 31 (1): 113–132.
- The Economist. 2004. *Inside the Mind of the Consumer*. June 12.
- Thompson, Clive. 2003. There's a Sucker Born in Every Medial Prefrontal Cortex. *New York Times Magazine*, October 25.
- Thompson, Jonathan. 2005. They Don't Just Want Your Money, They Want Your Brain. *London Independent on Sunday*, September 11.
- Tiltman, David. 2005. Mind Reading. *Marketing Research Bulletin*. <http://www.brandrepublic.com/bulletins/marketresearch/article/529442/market-research-mind-reading> (retrieved January 29, 2007).
- Toenjes, Richard. 2002. Why Be Moral in Business? A Rawlsian Approach to Moral Motivation. *Business Ethics Quarterly*, 12 (1): 57–72.
- Wallerstein, Immanuel. 1997. Social Science and the Quest for a Just Society. *American Journal of Sociology*, 102 (March): 1241–1257.
- Yoon, Carolyn, Angela H. Gutchess, Fred Feinberg, and Thad A. Polk. 2006. A Functional Magnetic Resonance Imaging Study of Neural Dissociations between Brand and Person Judgments. *Journal of Consumer Research*, 33 (1): 31–40.
- Young, Charles. 2002. Brain Waves, Picture Sorts, and Branding Moments. *Journal of Advertising Research*, 42 (4): 42–53.
- Zajonc, Robert B. 1998. Emotions. In *The Handbook of Social Psychology*, edited by Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey. 4th Edition. New York: McGraw-Hill Companies, Inc.
- Zak, Paul A. 2004. Neuroeconomics. *Philosophical Transactions of the Royal Society of London B*, 359: 1737–1748.

- Zak, Paul J., Karla Borja, William T. Matsner, and Robert Kurzban. 2005. The Neuroeconomics of Distrust: Sex Differences in Behavior and Physiology. *American Economics Association Papers and Proceedings*, 95 (2): 360–363.
- Zanetti, Veronique. 2001. Global Justice: Is Interventionism Desirable? *Metaphilosophy*, 32 (1/2): 196–211.
- Zinkhan, George M., Michael Bisessi, and Mary Jane Saxton. 1989. MBAs Changing Attitudes Toward Marketing Dilemmas: 1981–1989. *Journal of Business Ethics*, 8 (12): 963–974.