Psychology as a Biological Science

NOBA

Copyright

Contact Information:

Noba Project
www.nobaproject.com
info@nobaproject.com

Contents

About Noba & Acknowledgements 5

Psychology as Science **6**

1 Why Science? 7
Edward Diener

2 Conducting Psychology Research in the Real World 18
Matthias R. Mehl

3 Psychophysiological Methods in Neuroscience 36
Zachary Infantolino & Gregory A. Miller

4 History of Psychology 50
David B. Baker & Heather Sperry

Biological Basis of Behavior **67**

5 The Brain 68
Diane Beck & Evelina Tapia

6 The Nervous System 85
Aneeq Ahmad

7 Evolutionary Theories in Psychology 105
David M. Buss

8 Hormones & Behavior 120
Randy J. Nelson

9 Biochemistry of Love 144
Sue Carter & Stephen Porges

10 Epigenetics in Psychology 160
Ian Weaver

11 The Nature-Nurture Question 184
Eric Turkheimer

Sensation and Perception **197**

12 Vision 198
Simona Buetti & Alejandro Lleras

13 Hearing 214
Andrew J. Oxenham

14 Taste and Smell 230
Linda Bartoshuk & Derek Snyder

15 Touch and Pain 242
Guro E. Løseth, Dan-Mikael Ellingson & Siri Leknes

16 The Vestibular System 264
Dora Angelaki & J. David Dickman

17 Multi-Modal Perception 288
Lorin Lachs

18 Eyewitness Testimony and Memory Biases 308
Cara Laney & Elizabeth F. Loftus

Learning and Memory 322

19 Factors Influencing Learning 323
Aaron Benjamin

20 Memory (Encoding, Storage, Retrieval) 337
Kathleen B. McDermott & Henry L. Roediger

21 Conditioning and Learning 361
Mark E. Bouton

22 Forgetting and Amnesia 387
Nicole Dudukovic & Brice Kuhl

Cognition and Language 399

23 Attention 400
Frances Friedrich

24 Intelligence 415
Robert Biswas-Diener

25 Judgment and Decision Making 430
Max H. Bazerman

Development 445

26 Cognitive Development in Childhood 446
Robert Siegler

27 Adolescent Development 461
Jennifer Lansford

28 Aging 474
Tara Queen & Jacqui Smith

Social **491**

29 Social Neuroscience 492
Tiffany A. Ito & Jennifer T. Kubota

30 Aggression and Violence 510
Brad J. Bushman

31 Helping and Prosocial Behavior 526
Dennis L. Poepsel & David A. Schroeder

Personality **543**

32 Personality Traits 544
Edward Diener & Richard E. Lucas

Emotions and Motivation **562**

33 Affective Neuroscience 563
Eddie Harmon-Jones & Cindy Harmon-Jones

34 Functions of Emotions 577
Hyisung Hwang & David Matsumoto

35 Drive States 592
Sudeep Bhatia & George Loewenstein

Psychological Disorders **605**

36 History of Mental Illness 606
Ingrid G. Farreras

37 Mood Disorders 619
Anda Gershon & Renee Thompson

38 Autism: Insights from the Study of the Social Brain 642
Kevin A. Pelphrey

39 Psychopharmacology 655
Susan Barron

Well-Being **669**

40 The Healthy Life 670
Emily Hooker & Sarah Pressman

Index 691

About Noba

The Diener Education Fund (DEF) is a non-profit organization founded with the mission of re-inventing higher education to serve the changing needs of students and professors. The initial focus of the DEF is on making information, especially of the type found in textbooks, widely available to people of all backgrounds. This mission is embodied in the Noba project.

Noba is an open and free online platform that provides high-quality, flexibly structured textbooks and educational materials. The goals of Noba are three-fold:

- To reduce financial burden on students by providing access to free educational content
- To provide instructors with a platform to customize educational content to better suit their curriculum
- To present material written by a collection of experts and authorities in the field

The Diener Education Fund was co-founded by Drs. Ed and Carol Diener. Ed was a professor emeritus at the University of Illinois, Urbana Champaign, and a professor at University of Virginia and the University of Utah, and a senior scientist at the Gallup Organization but passed away in April 2021. For more information, please see **http://noba.to/78vdj2x5**. Carol Diener is the former director of the Mental Health Worker and the Juvenile Justice Programs at the University of Illinois. Both Ed and Carol are award- winning university teachers.

Acknowledgements

The Diener Education Fund would like to acknowledge the following individuals and companies for their contribution to the Noba Project: Robert Biswas-Diener as Managing Editor, Peter Lindberg as the former Operations Manager, and Nadezhda Lyubchik as the current Operations Manager; The Other Firm for user experience design and web development; Sockeye Creative for their work on brand and identity development; Arthur Mount for illustrations; Chad Hurst for photography; EEI Communications for manuscript proofreading; Marissa Diener, Shigehiro Oishi, Daniel Simons, Robert Levine, Lorin Lachs and Thomas Sander for their feedback and suggestions in the early stages of the project.

Psychology as Science

1
Why Science?

Edward Diener

Scientific research has been one of the great drivers of progress in human history, and the dramatic changes we have seen during the past century are due primarily to scientific findings—modern medicine, electronics, automobiles and jets, birth control, and a host of other helpful inventions. Psychologists believe that scientific methods can be used in the behavioral domain to understand and improve the world. Although psychology trails the biological and physical sciences in terms of progress, we are optimistic based on discoveries to date that scientific psychology will make many important discoveries that can benefit humanity. This module outlines the characteristics of the science, and the promises it holds for understanding behavior. The ethics that guide psychological research are briefly described. It concludes with the reasons you should learn about scientific psychology

Learning Objectives

- Describe how scientific research has changed the world.
- Describe the key characteristics of the scientific approach.
- Discuss a few of the benefits, as well as problems that have been created by science.
- Describe several ways that psychological science has improved the world.
- Describe a number of the ethical guidelines that psychologists follow.

Scientific Advances and World Progress

There are many people who have made positive contributions to humanity in modern times.

Take a careful look at the names on the following list. Which of these individuals do you think has helped humanity the most?

1. Mother Teresa
2. Albert Schweitzer
3. Edward Jenner
4. Norman Borlaug
5. Fritz Haber

The usual response to this question is "Who on earth are Jenner, Borlaug, and Haber?" Many people know that Mother Teresa helped thousands of people living in the slums of Kolkata (Calcutta). Others recall that Albert Schweitzer opened his famous hospital in Africa and went on to earn the Nobel Peace Prize. The other three historical figures, on the other hand, are far less well known. Jenner, Borlaug, and Haber were scientists whose research discoveries saved millions, and even billions, of lives. Dr. Edward Jenner is often considered the "father of immunology" because he was among the first to conceive of and test vaccinations. His pioneering work led directly to the eradication of smallpox. Many other diseases have been greatly reduced because of vaccines discovered using science—measles, pertussis, diphtheria, tetanus, typhoid, cholera, polio, hepatitis—and all are the legacy of Jenner. Fritz Haber and Norman Borlaug saved more than a billion human lives. They created the "Green Revolution" by producing hybrid agricultural crops and synthetic fertilizer. Humanity can now produce food for the seven billion people on the planet, and the starvation that does occur is related to political and economic factors rather than our collective ability to produce food.

Due to the breakthrough work of Dr. Edward Jenner, millions of vaccinations are now administered around the world every year preventing the spread of many treatable diseases while saving the lives of people of all ages. [Image: CDC Global Health, https://goo.gl/hokiWz, CC BY 2.0, https://goo.gl/9uSnqN]

If you examine major social and technological changes over the past century most of them can be directly attributed to science. The world in 1914 was very different than the one we see today (Easterbrook, 2003). There were few cars and most people traveled by foot, horseback, or carriage. There were no radios, televisions, birth control pills, artificial hearts or antibiotics. Only a small portion of the world had telephones, refrigeration or electricity. These days we find that 80% of all households have television and 84% have electricity. It is estimated that three quarters of the world's population has access to a mobile phone! Life expectancy was 47 years in 1900 and 79 years in 2010. The percentage of hungry and malnourished people in the world has dropped substantially across the globe. Even average levels of I.Q. have risen dramatically over the past century due to better nutrition and schooling.

All of these medical advances and technological innovations are the direct result of scientific research and understanding. In the modern age it is easy to grow complacent about the advances of science but make no mistake about it—science has made fantastic discoveries, and continues to do so. These discoveries have completely changed our world.

What Is Science?

What is this process we call "science," which has so dramatically changed the world? Ancient people were more likely to believe in magical and supernatural explanations for natural phenomena such as solar eclipses or thunderstorms. By contrast, scientifically minded people try to figure out the natural world through testing and observation. Specifically, science is the use of **systematic observation** in order to acquire knowledge. For example, children in a science class might combine vinegar and baking soda to observe the bubbly chemical reaction. These **empirical methods** are wonderful ways to learn about the physical and biological world. Science is not magic—it will not solve all human problems, and might not answer all our questions about behavior. Nevertheless, it appears to be the most powerful method we have for acquiring knowledge about the observable world. The essential elements of science are as follows:

1. *Systematic observation is the core of science*. Scientists observe the world, in a very organized way. We often measure the phenomenon we are observing. We record our observations so that memory biases are less likely to enter in to our conclusions. We are systematic in that we try to observe under controlled conditions, and also systematically vary the conditions of our observations so that we can see variations in the phenomena and understand when they occur and do not occur.
2. *Observation leads to hypotheses we can test*. When we develop **hypotheses**and **theories**, we state them in a way that can be tested. For example, you might make the claim that candles made of paraffin wax burn more slowly than do candles of the exact same size and shape made from bee's wax. This claim can be readily tested by timing the burning speed of

candles made from these materials.

3. *Science is democratic*. People in ancient times may have been willing to accept the views of their kings or pharaohs as absolute truth. These days, however, people are more likely to want to be able to form their own opinions and debate conclusions. Scientists are skeptical and have open discussions about their observations and theories. These debates often occur as scientists publish competing findings with the idea that the best data will win the argument.

Systematic observation is the core of science. [Image: Cvl Neuro, https://goo.gl/Avbju7, CC BY-SA 3.0, https://goo.gl/uhHola]

4. *Science is cumulative*. We can learn the important truths discovered by earlier scientists and build on them. Any physics student today knows more about physics than Sir Isaac Newton did even though Newton was possibly the most brilliant physicist of all time. A crucial aspect of scientific progress is that after we learn of earlier advances, we can build upon them and move farther along the path of knowledge.

Psychology as a Science

Even in modern times many people are skeptical that psychology is really a science. To some degree this doubt stems from the fact that many psychological phenomena such as depression, intelligence, and prejudice do not seem to be directly observable in the same way that we can observe the changes in ocean tides or the speed of light. Because thoughts and feelings are invisible many early psychological researchers chose to focus on behavior. You might have noticed that some people act in a friendly and outgoing way while others appear to be shy and withdrawn. If you have made these types of observations then you are acting just like early psychologists who used behavior to draw inferences about various types of personality. By using behavioral measures and rating scales it is possible to measure thoughts and feelings. This is similar to how other researchers explore “invisible” phenomena such as the way that educators measure academic performance or economists measure quality of life.

One important pioneering researcher was Francis Galton, a cousin of Charles Darwin who lived in England during the late 1800s. Galton used patches of color to test people’s ability to distinguish between them. He also invented the self-report questionnaire, in which people offered their own expressed judgments or opinions on various matters. Galton was able to use self-reports to examine—among other things—people’s differing ability to accurately judge distances.

In 1875 Francis Galton did pioneering studies of twins to determine how much the similarities and differences in twins were affected by their life experiences. In the course of this work he coined the phrase "Nature versus Nurture". [Image: XT Inc., https://goo.gl/F1Wvu7, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Although he lacked a modern understanding of genetics Galton also had the idea that scientists could look at the behaviors of identical and fraternal twins to estimate the degree to which genetic and social factors contribute to personality; a puzzling issue we currently refer to as the “nature-nurture question.”

In modern times psychology has become more sophisticated. Researchers now use better measures, more sophisticated study designs and better statistical analyses to explore human nature. Simply take the example of studying the emotion of happiness. How would you go about studying happiness? One straight-forward method is to simply ask people about their happiness and to have them use a numbered scale to indicate their feelings. There are, of course, several problems with this. People might lie about their happiness, might not be able to accurately report on their own happiness, or might not use the numerical scale in the same way. With these limitations in mind modern psychologists employ a wide range of methods to assess happiness. They use, for instance, “peer report measures” in which they ask close friends and family members about the happiness of a target individual. Researchers can then compare these ratings to the self-report ratings and check for discrepancies. Researchers also use memory measures, with the idea that dispositionally positive people have an easier time recalling pleasant events and negative people have an easier time recalling unpleasant events. Modern psychologists even use biological measures such as saliva cortisol samples (cortisol is a stress related hormone) or fMRI images of brain activation (the left pre-frontal cortex is one area of brain activity associated with good moods).

Despite our various methodological advances it is true that psychology is still a very young science. While physics and chemistry are hundreds of years old psychology is barely a hundred and fifty years old and most of our major findings have occurred only in the last 60 years. There are legitimate limits to psychological science but it is a science nonetheless.

Psychological Science is Useful

Psychological science is useful for creating interventions that help people live better lives. A growing body of research is concerned with determining which therapies are the most and least effective for the treatment of psychological disorders.

For example, many studies have shown that cognitive behavioral therapy can help many people suffering from depression and anxiety disorders (Butler, Chapman, Forman, & Beck, 2006; Hoffman & Smits, 2008). In contrast, research reveals that some types of therapies actually might be harmful on average (Lilienfeld, 2007).

Cognitive Behavioral Therapy has shown to be effective in treating a variety of conditions, including depression. [Image: SalFalco, https://goo.gl/3knLoJ, CC BY-NC 2.0, https://goo.gl/HEXbAA]

In organizational psychology, a number of psychological interventions have been found by researchers to produce greater productivity and satisfaction in the workplace (e.g., Guzzo, Jette, & Katzell, 1985). Human factor engineers have greatly increased the safety and utility of the products we use. For example, the human factors psychologist Alphonse Chapanis and other researchers redesigned the cockpit controls of aircraft to make them less confusing and easier to respond to, and this led to a decrease in pilot errors and crashes.

Forensic sciences have made courtroom decisions more valid. We all know of the famous cases of imprisoned persons who have been exonerated because of DNA evidence. Equally dramatic cases hinge on psychological findings. For instance, psychologist Elizabeth Loftus has conducted research demonstrating the limits and unreliability of eyewitness testimony and memory. Thus, psychological findings are having practical importance in the world outside the laboratory. Psychological science has experienced enough success to demonstrate that it works, but there remains a huge amount yet to be learned.

Ethics of Scientific Psychology

Psychology differs somewhat from the natural sciences such as chemistry in that researchers conduct studies with human research participants. Because of this there is a natural tendency to want to guard research participants against potential psychological harm. For example, it might be interesting to see how people handle ridicule but it might not be advisable to ridicule research participants.

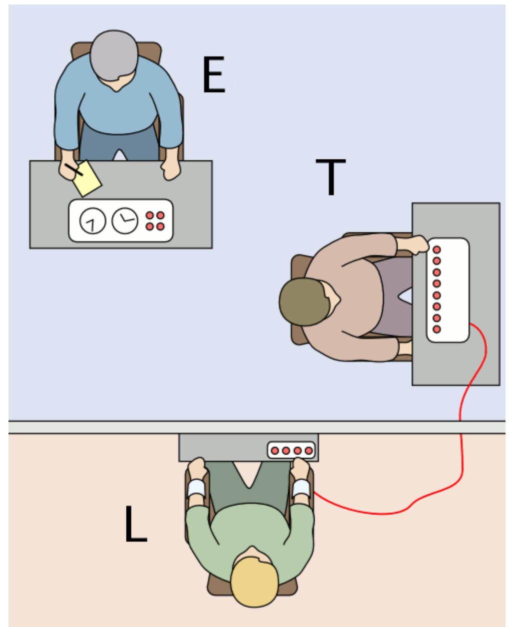


Diagram of the Milgram Experiment in which the "teacher" (T) was asked to deliver a (supposedly) painful electric shock to the "learner"(L). Would this experiment be approved by a review board today? [Image: Fred the Oyster, https://goo.gl/ZIbQz1, CC BY-SA 4.0, https://goo.gl/X3i0tq]

Scientific psychologists follow a specific set of guidelines for research known as a code of ethics. There are extensive ethical guidelines for how human participants should be treated in psychological research (Diener & Crandall, 1978; Sales & Folkman, 2000). Following are a few highlights:

1. *Informed consent*. In general, people should know when they are involved in research, and understand what will happen to them during the study. They should then be given a free choice as to whether to participate.
2. *Confidentiality*. Information that researchers learn about individual participants should not be made public without the consent of the individual.
3. *Privacy*. Researchers should not make observations of people in private places such as their bedrooms without their knowledge and consent. Researchers should not seek confidential information from others, such as school authorities, without consent of the participant or his or her guardian.
4. *Benefits*. Researchers should consider the benefits of their proposed research and weigh these against potential risks to the participants. People who participate in psychological studies should be exposed to risk only if they fully understand these risks and only if the likely benefits clearly outweigh the risks.
5. *Deception*. Some researchers need to deceive participants in order to hide the true nature of the study. This is typically done to prevent participants from modifying their behavior

in unnatural ways. Researchers are required to “debrief” their participants after they have completed the study. Debriefing is an opportunity to educate participants about the true nature of the study.

Why Learn About Scientific Psychology?

I once had a psychology professor who asked my class why we were taking a psychology course. Our responses give the range of reasons that people want to learn about psychology:

1. To understand ourselves
2. To understand other people and groups
3. To be better able to influence others, for example, in socializing children or motivating employees
4. To learn how to better help others and improve the world, for example, by doing effective psychotherapy
5. To learn a skill that will lead to a profession such as being a social worker or a professor
6. To learn how to evaluate the research claims you hear or read about
7. Because it is interesting, challenging, and fun! People want to learn about psychology because this is exciting in itself, regardless of other positive outcomes it might have. Why do we see movies? Because they are fun and exciting, and we need no other reason. Thus, one good reason to study psychology is that it can be rewarding in itself.

Conclusions

The science of psychology is an exciting adventure. Whether you will become a scientific psychologist, an applied psychologist, or an educated person who knows about psychological research, this field can influence your life and provide fun, rewards, and understanding. My hope is that you learn a lot from the modules in this e-text, and also that you enjoy the experience! I love learning about psychology and neuroscience, and hope you will too!

Outside Resources

Web: Science Heroes- A celebration of people who have made lifesaving discoveries.
http://www.scienceheroes.com/index.php?option=com_content&view=article&id=258&Itemid=27

Discussion Questions

1. Some claim that science has done more harm than good. What do you think?
2. Humanity is faced with many challenges and problems. Which of these are due to human behavior, and which are external to human actions?
3. If you were a research psychologist, what phenomena or behaviors would most interest you?
4. Will psychological scientists be able to help with the current challenges humanity faces, such as global warming, war, inequality, and mental illness?
5. What can science study and what is outside the realm of science? What questions are impossible for scientists to study?
6. Some claim that science will replace religion by providing sound knowledge instead of myths to explain the world. They claim that science is a much more reliable source of solutions to problems such as disease than is religion. What do you think? Will science replace religion, and should it?
7. Are there human behaviors that should not be studied? Are some things so sacred or dangerous that we should not study them?

Vocabulary

Empirical methods
Approaches to inquiry that are tied to actual measurement and observation.

Ethics
Professional guidelines that offer researchers a template for making decisions that protect research participants from potential harm and that help steer scientists away from conflicts of interest or other situations that might compromise the integrity of their research.

Hypotheses
A logical idea that can be tested.

Systematic observation
The careful observation of the natural world with the aim of better understanding it. Observations provide the basic data that allow scientists to track, tally, or otherwise organize information about the natural world.

Theories
Groups of closely related phenomena or observations.

References

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review, 26*, 17–31.

Diener, E., & Crandall, R. (1978). *Ethics in social and behavioral research.* Chicago, IL: University of Chicago Press.

Easterbrook, G. (2003). *The progress paradox*. New York, NY: Random House.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. (1985). The effects of psychologically based intervention programs on worker productivity: A meta-analysis. *Personnel Psychology, 38*, 275.291.

Hoffman, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders. *Journal of Clinical Psychiatry, 69*, 621–32.

Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2*, 53–70.

Moore, D. (2003). Public lukewarm on animal rights. Gallup News Service, May 21. http://www.gallup.com/poll/8461/public-lukewarm-animal-rights.aspx

Sales, B. D., & Folkman, S. (Eds.). (2000). *Ethics in research with human participants.* Washington, DC: American Psychological Association.

2

Conducting Psychology Research in the Real World

Matthias R. Mehl

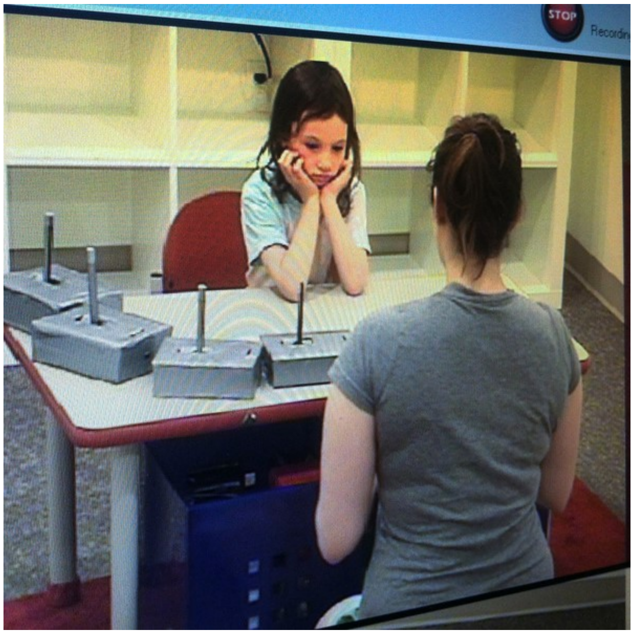
Because of its ability to determine cause-and-effect relationships, the laboratory experiment is traditionally considered the method of choice for psychological science. One downside, however, is that as it carefully controls conditions and their effects, it can yield findings that are out of touch with reality and have limited use when trying to understand real-world behavior. This module highlights the importance of also conducting research outside the psychology laboratory, within participants' natural, everyday environments, and reviews existing methodologies for studying daily life

Learning Objectives

- Identify limitations of the traditional laboratory experiment.
- Explain ways in which daily life research can further psychological science.
- Know what methods exist for conducting psychological research in the real world.

Introduction

The laboratory experiment is traditionally considered the "gold standard" in psychology research. This is because only laboratory experiments can clearly separate cause from effect and therefore establish causality. Despite this unique strength, it is also clear that a scientific field that is mainly based on controlled laboratory studies ends up lopsided. Specifically, it accumulates a lot of knowledge on what *can* happen—under carefully isolated and controlled circumstances—but it has little to say about what actually *does* happen under the circumstances that people actually encounter in their daily lives.



Do the research results obtained in isolated, carefully controlled laboratory conditions generalize into the real world? [Image: Nessen Marshall, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

For example, imagine you are a participant in an experiment that looks at the effect of being in a good mood on generosity, a topic that may have a good deal of practical application. Researchers create an internally-valid, carefully-controlled experiment where they randomly assign you to watch either a happy movie or a neutral movie, and then you are given the opportunity to help the researcher out by staying longer and participating in another study. If people in a good mood are more willing to stay and help out, the researchers can feel confident that – since everything else was held constant – your positive mood led you to be more helpful. However, what does this tell us about helping behaviors in the real world? Does it generalize to other kinds of helping, such as donating money to a charitable cause? Would all kinds of happy movies produce this behavior, or only this one? What about other positive experiences that might boost mood, like receiving a compliment or a good grade? And what if you were watching the movie with friends, in a crowded theatre, rather than in a sterile research lab? Taking research out into the real world can help answer some of these sorts of important questions.

As one of the founding fathers of social psychology remarked, “Experimentation in the laboratory occurs, socially speaking, on an island quite isolated from the life of society” (Lewin, 1944, p. 286). This module highlights the importance of going beyond experimentation and also conducting research outside the laboratory (Reis & Gosling, 2010), directly within participants’ natural environments, and reviews existing methodologies for studying daily life.

Rationale for Conducting Psychology Research in the Real World

One important challenge researchers face when designing a study is to find the right balance between ensuring **internal validity**, or the degree to which a study allows unambiguous causal inferences, and **external validity**, or the degree to which a study ensures that potential findings apply to settings and samples other than the ones being studied (Brewer, 2000). Unfortunately, these two kinds of validity tend to be difficult to achieve at the same time, in one study. This is because creating a controlled setting, in which all potentially influential factors (other than the experimentally-manipulated variable) are controlled, is bound to create an environment that is quite different from what people naturally encounter (e.g., using a happy movie clip to promote helpful behavior). However, it is the degree to which an experimental situation is comparable to the corresponding real-world situation of interest that determines how generalizable potential findings will be. In other words, if an experiment is very far-off from what a person might normally experience in everyday life, you might reasonably question just how useful its findings are.

Because of the incompatibility of the two types of validity, one is often—by design—prioritized over the other. Due to the importance of identifying true causal relationships, psychology has traditionally emphasized internal over external validity. However, in order to make claims about human behavior that apply across populations and environments, researchers complement traditional laboratory research, where participants are brought into the lab, with field research where, in essence, the psychological laboratory is brought to participants. Field studies allow for the important test of how psychological variables and processes of interest "behave" under real-world circumstances (i.e., what *actually does happen* rather than what *can happen*). They can also facilitate "downstream" operationalizations of constructs that measure life outcomes of interest *directly* rather than indirectly.

Take, for example, the fascinating field of psychoneuroimmunology, where the goal is to understand the interplay of psychological factors - such as personality traits or one's stress level - and the immune system. Highly sophisticated and carefully controlled experiments offer ways to isolate the variety of neural, hormonal, and cellular mechanisms that link psychological variables such as chronic stress to biological outcomes such as immunosuppression (a state of impaired immune functioning; Sapolsky, 2004). Although these studies demonstrate impressively how psychological factors can affect health-relevant biological processes, they—because of their research design—remain mute about the degree to which these factors actually do undermine people's everyday health in real life. It is certainly important to show that laboratory stress can alter the number of natural killer cells in the blood. But it is equally important to test to what extent the levels of stress that people experience on a day-to-day basis result in them catching a cold more often or taking longer to recover from one. The goal for researchers, therefore, must be to complement traditional laboratory experiments with less controlled studies under real-world circumstances. The term **ecological validity** is used to refer the degree to which an effect has been obtained under conditions that are typical for what happens in everyday life (Brewer, 2000). In this example, then, people might keep a careful daily log of how much stress they are under as well as noting physical symptoms such as headaches or nausea. Although many factors beyond stress level may be responsible for these symptoms, this more correlational approach can shed light on how the relationship between stress and health plays out outside of the laboratory.

An Overview of Research Methods for Studying Daily Life

Capturing "life as it is lived" has been a strong goal for some researchers for a long time. Wilhelm and his colleagues recently published a comprehensive review of early attempts to systematically document daily life (Wilhelm, Perrez, & Pawlik, 2012). Building onto these original methods, researchers have, over the past decades, developed a broad toolbox for measuring experiences, behavior, and physiology directly in participants' daily lives (Mehl & Conner, 2012). Figure 1 provides a schematic overview of the methodologies described below.

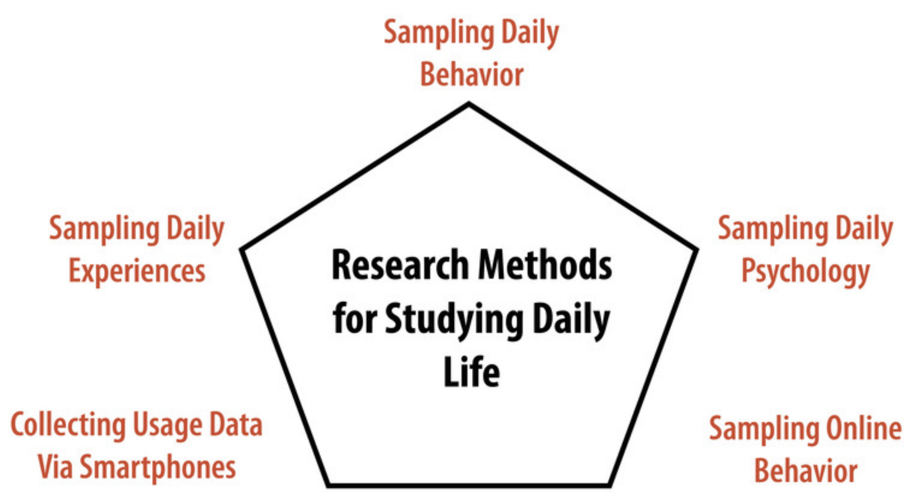


Figure 1. Schematic Overview of Research Methods for Studying Daily Life

Studying Daily Experiences

Starting in the mid-1970s, motivated by a growing skepticism toward highly-controlled laboratory studies, a few groups of researchers developed a set of new methods that are now commonly known as the experience-sampling method (Hektner, Schmidt, & Csikszentmihalyi, 2007), ecological momentary assessment (Stone & Shiffman, 1994), or the

diary method (Bolger & Rafaeli, 2003). Although variations within this set of methods exist, the basic idea behind all of them is to collect in-the-moment (or, close-to-the-moment) self-report data directly from people as they go about their daily lives. This is typically accomplished by asking participants' repeatedly (e.g., five times per day) over a period of time (e.g., a week) to report on their current thoughts and feelings. The momentary questionnaires often ask about their location (e.g., "Where are you now?"), social environment (e.g., "With whom are you now?"), activity (e.g., "What are you currently doing?"), and experiences (e.g., "How are you feeling?"). That way, researchers get a snapshot of what was going on in participants' lives at the time at which they were asked to report.

Technology has made this sort of research possible, and recent technological advances have altered the different tools researchers are able to easily use. Initially, participants wore electronic wristwatches that beeped at preprogrammed but seemingly random times, at which they completed one of a stack of provided paper questionnaires. With the mobile computing revolution, both the prompting and the questionnaire completion were gradually replaced by handheld devices such as smartphones. Being able to collect the momentary questionnaires digitally and time-stamped (i.e., having a record of exactly when participants responded) had major methodological and practical advantages and contributed to experience sampling going mainstream (Conner, Tennen, Fleeson, & Barrett, 2009).

Using modern technology like smartphones allows for more widespread experience sampling of research participants. Whether at home, work, or just sitting in a coffee shop technology makes it easier than ever to participate in psychology research. [Image: Vladimir Yaitskiy, https://goo.gl/7sjXfq, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Over time, experience sampling and related momentary self-report methods have become very popular, and, by now, they are effectively the gold standard for studying daily life. They have helped make progress in almost all areas of psychology (Mehl & Conner, 2012). These methods ensure receiving many measurements from many participants, and has further inspired the development of novel statistical methods (Bolger & Laurenceau, 2013). Finally, and maybe most importantly, they accomplished what they sought out to accomplish: to bring attention to what psychology ultimately wants and needs to know about, namely "what people actually do, think, and feel in the various contexts of their lives" (Funder, 2001, p. 213). In short, these approaches have allowed researchers to do research that is more externally valid, or more generalizable to real life, than the traditional laboratory experiment.

To illustrate these techniques, consider a classic study, Stone, Reed, and Neale (1987), who tracked positive and negative experiences surrounding a respiratory infection using daily experience sampling. They found that undesirable experiences peaked and desirable ones dipped about four to five days prior to participants coming down with the cold. More recently, Killingsworth and Gilbert (2010) collected momentary self-reports from more than 2,000 participants via a smartphone app. They found that participants were less happy when their mind was in an idling, mind-wandering state, such as surfing the Internet or multitasking at work, than when it was in an engaged, task-focused one, such as working diligently on a paper. These are just two examples that illustrate how experience-sampling studies have yielded findings that could not be obtained with traditional laboratory methods.

Recently, the **day reconstruction method (DRM)** (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004) has been developed to obtain information about a person's daily experiences without going through the burden of collecting momentary experience-sampling data. In the DRM, participants report their experiences of a given day retrospectively after engaging in a systematic, experiential reconstruction of the day on the following day. As a participant in this type of study, you might look back on yesterday, divide it up into a series of episodes such as "made breakfast," "drove to work," "had a meeting," etc. You might then report who you were with in each episode and how you felt in each. This approach has shed light on what situations lead to moments of positive and negative mood throughout the course of a normal day.

Studying Daily Behavior

Experience sampling is often used to study everyday behavior (i.e., daily social interactions and activities). In the laboratory, behavior is best studied using direct behavioral observation (e.g., video recordings). In the real world, this is, of course, much more difficult. As Funder put it, it seems it would require a "detective's report [that] would specify in exact detail everything the participant said and did, and with whom, in all of the contexts of the participant's life" (Funder, 2007, p. 41).

As difficult as this may seem, Mehl and colleagues have developed a naturalistic observation methodology that is similar in spirit. Rather than following participants—like a detective—with a video camera (see Craik, 2000), they equip participants with a portable audio recorder that is programmed to periodically record brief snippets of ambient sounds (e.g., 30 seconds every 12 minutes). Participants carry the recorder (originally a microcassette recorder, now a smartphone app) on them as they go about their days and return it at the end of the study. The recorder provides researchers with a series of sound bites that, together, amount to an acoustic diary of participants' days as they naturally unfold—and that constitute a representative sample of their daily activities and social encounters. Because it is somewhat similar to having the researcher's ear at the participant's lapel, they called their method the **electronically activated recorder, or EAR** (Mehl, Pennebaker, Crow, Dabbs, & Price, 2001). The ambient sound recordings can be coded for many things, including participants' locations (e.g., at school, in a coffee shop), activities (e.g., watching TV, eating), interactions (e.g., in a group, on the phone), and emotional expressions (e.g., laughing, sighing). As unnatural or intrusive as it might seem, participants report that they quickly grow accustomed to the EAR and say they soon find themselves behaving as they normally would.

In a cross-cultural study, Ramírez-Esparza and her colleagues used the EAR method to study sociability in the United States and Mexico. Interestingly, they found that although American participants rated themselves significantly higher than Mexicans on the question, "I see myself as a person who is talkative," they actually spent almost 10 percent less time talking than Mexicans did (Ramírez-Esparza, Mehl, Álvarez Bermúdez, & Pennebaker, 2009). In a similar way, Mehl and his colleagues used the EAR method to debunk the long-standing myth that women are considerably more talkative than men. Using data from six different studies, they showed that both sexes use on average about 16,000 words per day. The estimated sex difference of 546 words was trivial compared to the immense range of more than 46,000 words between the least and most talkative individual (695 versus 47,016 words; Mehl, Vazire, Ramírez-Esparza, Slatcher, & Pennebaker, 2007). Together, these studies demonstrate how naturalistic observation can be used to study objective aspects of daily behavior and how it can yield findings quite different from what other methods yield (Mehl, Robbins, & Deters, 2012).

A series of other methods and creative ways for assessing behavior directly and unobtrusively in the real world are described in a seminal book on real-world, subtle measures (Webb, Campbell, Schwartz, Sechrest, & Grove, 1981). For example, researchers have used time-lapse photography to study the flow of people and the use of space in urban public places (Whyte, 1980). More recently, they have observed people's personal (e.g., dorm rooms) and professional (e.g., offices) spaces to understand how personality is expressed and detected in everyday environments (Gosling, Ko, Mannarelli, & Morris, 2002). They have even systematically collected and analyzed people's garbage to measure what people actually consume (e.g., empty alcohol bottles or cigarette boxes) rather than what they *say* they consume (Rathje & Murphy, 2001). Because people often cannot and sometimes may not want to accurately report what they do, the direct—and ideally nonreactive—assessment of real-world behavior is of high importance for psychological research (Baumeister, Vohs, &

Funder, 2007).

Studying Daily Physiology

In addition to studying how people think, feel, and behave in the real world, researchers are also interested in how our bodies respond to the fluctuating demands of our lives. What are the daily experiences that make our "blood boil"? How do our neurotransmitters and hormones respond to the stressors we encounter in our lives? What physiological reactions do we show to being loved—or getting ostracized? You can see how studying these powerful experiences in real life, as they actually happen, may provide more rich and informative data than one might obtain in an artificial laboratory setting that merely mimics these experiences.

Real world stressors may result in very different physiological responses than the same stressors simulated in a lab environment. [Image: State Farm, https://goo.gl/FGYyVz, CC BY 2.0, https://goo.gl/9uSnqN]

Also, in pursuing these questions, it is important to keep in mind that what is stressful, engaging, or boring for one person might not be so for another. It is, in part, for this reason that researchers have found only limited correspondence between how people respond physiologically to a standardized laboratory stressor (e.g., giving a speech) and how they respond to stressful experiences in their lives. To give an example, Wilhelm and Grossman (2010) describe a participant who showed rather minimal heart rate increases in response to a laboratory stressor (about five to 10 beats per minute) but quite dramatic increases (almost 50 beats per minute) later in the afternoon while watching a soccer game. Of course, the reverse pattern can happen as well, such as when patients have high blood pressure in the doctor's office but not in their home environment—the so-called white coat hypertension (White, Schulman, McCabe, & Dey, 1989).

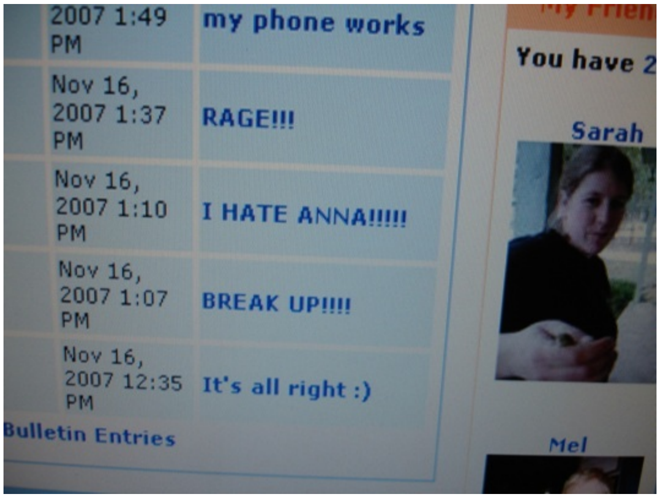
Ambulatory physiological monitoring – that is, monitoring physiological reactions as people go about their daily lives - has a long history in biomedical research and an array of monitoring devices exist (Fahrenberg & Myrtek, 1996). Among the biological signals that can now be measured in daily life with portable signal recording devices are the electrocardiogram (ECG), blood pressure, electrodermal activity (or "sweat response"), body temperature, and even the electroencephalogram (EEG) (Wilhelm & Grossman, 2010). Most recently, researchers have added **ambulatory assessment** of hormones (e.g., cortisol) and other biomarkers (e.g., immune markers) to the list (Schlotz, 2012). The development of ever more sophisticated ways to track what goes on underneath our skins as we go about our lives is a fascinating and rapidly advancing field.

In a recent study, Lane, Zareba, Reis, Peterson, and Moss (2011) used experience sampling combined with ambulatory electrocardiography (a so-called Holter monitor) to study how emotional experiences can alter cardiac function in patients with a congenital heart abnormality (e.g., long QT syndrome). Consistent with the idea that emotions may, in some cases, be able to trigger a cardiac event, they found that typical—in most cases even relatively low intensity— daily emotions had a measurable effect on ventricular repolarization, an important cardiac indicator that, in these patients, is linked to risk of a cardiac event. In another study, Smyth and colleagues (1998) combined experience sampling with momentary assessment of cortisol, a stress hormone. They found that momentary reports of current or even anticipated stress predicted increased cortisol secretion 20 minutes later. Further, and independent of that, the experience of other kinds of negative affect (e.g., anger, frustration) also predicted higher levels of cortisol and the experience of positive affect (e.g., happy, joyful) predicted lower levels of this important stress hormone. Taken together, these studies illustrate how researchers can use ambulatory physiological monitoring to study how the little —and seemingly trivial or inconsequential—experiences in our lives leave objective, measurable traces in our bodily systems.

Studying Online Behavior

Another domain of daily life that has only recently emerged is virtual daily behavior or how people act and interact with others on the Internet. Irrespective of whether social media will turn out to be humanity's blessing or curse (both scientists and laypeople are currently divided over this question), the fact is that people are spending an ever increasing amount of time online. In light of that, researchers are beginning to think of virtual behavior as being as serious as "actual" behavior and seek to make it a legitimate target of their investigations (Gosling & Johnson, 2010).

One way to study virtual behavior is to make use of the fact that most of what people do on the Web—emailing, chatting, tweeting, blogging, posting— leaves direct (and permanent) verbal traces. For example, differences in the ways in which people use words (e.g., subtle



Online activity reveals a lot of psychological information to researchers. [Image: Sarah C. Frey, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

preferences in word choice) have been found to carry a lot of psychological information (Pennebaker, Mehl, & Niederhoffer, 2003). Therefore, a good way to study virtual social behavior is to study virtual language behavior. Researchers can download people's—often public—verbal expressions and communications and analyze them using modern text analysis programs (e.g., Pennebaker, Booth, & Francis, 2007).

For example, Cohn, Mehl, and Pennebaker (2004) downloaded blogs of more than a thousand users of lifejournal.com, one of the first Internet blogging sites, to study how people responded socially and emotionally to the attacks of September 11, 2001. In going "the online route," they could bypass a critical limitation of coping research, the inability to obtain baseline information; that is, how people were doing *before* the traumatic event occurred. Through access to the database of public blogs, they downloaded entries from two months prior to two months after the attacks. Their linguistic analyses revealed that in the first days after the attacks, participants expectedly expressed more negative emotions and were more cognitively and socially engaged, asking questions and sending messages of support. Already after two weeks, though, their moods and social engagement returned to baseline, and, interestingly, their use of cognitive-analytic words (e.g., "think," "question") even dropped below their normal level. Over the next six weeks, their mood hovered around their pre-9/11 baseline, but both their social engagement and cognitive-analytic processing stayed remarkably low. This suggests a social and cognitive weariness in the aftermath of the attacks. In using virtual verbal behavior as a marker of psychological functioning, this study was able to draw a fine timeline of how humans cope with disasters.

Reflecting their rapidly growing real-world importance, researchers are now beginning to investigate behavior on social networking sites such as Facebook (Wilson, Gosling, & Graham, 2012). Most research looks at psychological correlates of online behavior such as personality traits and the quality of one's social life but, importantly, there are also first attempts to export traditional experimental research designs into an online setting. In a pioneering study of online social influence, Bond and colleagues (2012) experimentally tested the effects that peer feedback has on voting behavior. Remarkably, their sample consisted of 16 million (!) Facebook users. They found that online political-mobilization messages (e.g., "I voted" accompanied by selected pictures of their Facebook friends) influenced real-world voting behavior. This was true not just for users who saw the messages but also for their friends and friends of their friends. Although the intervention effect on a single user was very small, through the enormous number of users and indirect social contagion effects, it resulted cumulatively in an estimated 340,000 additional votes—enough to tilt a close election. In short, although still in its infancy, research on virtual daily behavior is bound to change social science, and it has already helped us better understand both virtual and "actual" behavior.

"Smartphone Psychology" ?

A review of research methods for studying daily life would not be complete without a vision of "what's next." Given how common they have become, it is safe to predict that smartphones will not just remain devices for everyday online communication but will also become devices for scientific data collection and intervention (Kaplan & Stone, 2013; Yarkoni, 2012). These devices automatically store vast amounts of real-world user interaction data, and, in addition, they are equipped with sensors to track the physical (e. g., location, position) and social (e.g., wireless connections around the phone) context of these interactions. Miller (2012, p. 234) states, "The question is not whether smartphones will revolutionize psychology but how, when, and where the revolution will happen." Obviously, their immense potential for data collection also brings with it big new challenges for researchers (e.g., privacy protection, data analysis, and synthesis). Yet it is clear that many of the methods described in this module—and many still to be developed ways of collecting real-world data—will, in the future, become integrated into the devices that people naturally and happily carry with them from the moment they get up in the morning to the moment they go to bed.

Conclusion

This module sought to make a case for psychology research conducted outside the lab. If the ultimate goal of the social and behavioral sciences is to explain human behavior, then researchers must also—in addition to conducting carefully controlled lab studies—deal with the "messy" real world and find ways to capture life as it naturally happens.

Mortensen and Cialdini (2010) refer to the dynamic give-and-take between laboratory and field research as "**full-cycle psychology**". Going full cycle, they suggest, means that "researchers use naturalistic observation to determine an effect's presence in the real world, theory to determine what processes underlie the effect, experimentation to verify the effect and its underlying processes, and a return to the natural environment to corroborate the experimental findings" (Mortensen & Cialdini, 2010, p. 53). To accomplish this, researchers have access to a toolbox of research methods for studying daily life that is now more diverse and more versatile than it has ever been before. So, all it takes is to go ahead and—literally—bring science to life.

Outside Resources

Website: Society for Ambulatory Assessment
http://www.ambulatory-assessment.org

Discussion Questions

1. What do you think about the tradeoff between unambiguously establishing cause and effect (internal validity) and ensuring that research findings apply to people's everyday lives (external validity)? Which one of these would you prioritize as a researcher? Why?
2. What challenges do you see that daily-life researchers may face in their studies? How can they be overcome?
3. What ethical issues can come up in daily-life studies? How can (or should) they be addressed?
4. How do you think smartphones and other mobile electronic devices will change psychological research? What are their promises for the field? And what are their pitfalls?

Vocabulary

Ambulatory assessment
An overarching term to describe methodologies that assess the behavior, physiology, experience, and environments of humans in naturalistic settings.

Daily Diary method
A methodology where participants complete a questionnaire about their thoughts, feelings, and behavior of the day at the end of the day.

Day reconstruction method (DRM)
A methodology where participants describe their experiences and behavior of a given day retrospectively upon a systematic reconstruction on the following day.

Ecological momentary assessment
An overarching term to describe methodologies that repeatedly sample participants' real-world experiences, behavior, and physiology in real time.

Ecological validity
The degree to which a study finding has been obtained under conditions that are typical for what happens in everyday life.

Electronically activated recorder, or EAR
A methodology where participants wear a small, portable audio recorder that intermittently records snippets of ambient sounds around them.

Experience-sampling method
A methodology where participants report on their momentary thoughts, feelings, and behaviors at different points in time over the course of a day.

External validity
The degree to which a finding generalizes from the specific sample and context of a study to some larger population and broader settings.

Full-cycle psychology
A scientific approach whereby researchers start with an observational field study to identify an effect in the real world, follow up with laboratory experimentation to verify the effect and isolate the causal mechanisms, and return to field research to corroborate their experimental

findings.

Generalize
Generalizing, in science, refers to the ability to arrive at broad conclusions based on a smaller sample of observations. For these conclusions to be true the sample should accurately represent the larger population from which it is drawn.

Internal validity
The degree to which a cause-effect relationship between two variables has been unambiguously established.

Linguistic inquiry and word count
A quantitative text analysis methodology that automatically extracts grammatical and psychological information from a text by counting word frequencies.

Lived day analysis
A methodology where a research team follows an individual around with a video camera to objectively document a person's daily life as it is lived.

White coat hypertension
A phenomenon in which patients exhibit elevated blood pressure in the hospital or doctor's office but not in their everyday lives.

References

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*, 396–403.

Bolger, N., & Laurenceau, J-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. Annual Review of Psychology, 54, 579–616.

Bond, R. M., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61 million-person experiment in social influence and political mobilization. *Nature*, 489, 295–298.

Brewer, M. B. (2000). Research design and issues of validity. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 3–16). New York, NY: Cambridge University Press.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic indicators of psychological change after September 11, 2001. Psychological Science, 15, 687–693.

Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and Personality Psychology Compass, 3*, 292–313.

Craik, K. H. (2000). The lived day of an individual: A person-environment perspective. In W. B. Walsh, K. H. Craik, & R. H. Price (Eds.), *Person-environment psychology: New directions and perspectives* (pp. 233–266). Mahwah, NJ: Lawrence Erlbaum Associates.

Fahrenberg, J., &. Myrtek, M. (Eds.) (1996). *Ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Seattle, WA: Hogrefe & Huber.

Funder, D. C. (2007). *The personality puzzle*. New York, NY: W. W. Norton & Co.

Funder, D. C. (2001). Personality. Review of Psychology, 52, 197–221.

Gosling, S. D., & Johnson, J. A. (2010). *Advanced methods for conducting online behavioral research*. Washington, DC: American Psychological Association.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*, 379–398.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage.

Kahneman, D., Krueger, A., Schkade, D., Schwarz, N., and Stone, A. (2004). A survey method for characterizing daily life experience: The Day Reconstruction Method. *Science*, 306, 1776–780.

Kaplan, R. M., & Stone A. A. (2013). Bringing the laboratory and clinic to the community: Mobile technologies for health promotion and disease prevention. *Annual Review of Psychology, 64*, 471-498.

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330, 932.

Lane, R. D., Zareba, W., Reis, H., Peterson, D., &, Moss, A. (2011). Changes in ventricular repolarization duration during typical daily emotion in patients with Long QT Syndrome. *Psychosomatic Medicine, 73*, 98–105.

Lewin, K. (1944) *Constructs in psychology and psychological ecology*. University of Iowa Studies in Child Welfare, 20, 23–27.

Mehl, M. R., & Conner, T. S. (Eds.) (2012). *Handbook of research methods for studying daily life*. New York, NY: Guilford Press.

Mehl, M. R., Pennebaker, J. W., Crow, M., Dabbs, J., & Price, J. (2001). The electronically activated recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers, 33*, 517–523.

Mehl, M. R., Robbins, M. L., & Deters, G. F. (2012). Naturalistic observation of health-relevant social processes: The electronically activated recorder (EAR) methodology in psychosomatics. *Psychosomatic Medicine, 74*, 410–417.

Mehl, M. R., Vazire, S., Ramírez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science, 317*, 82.

Miller, G. (2012). The smartphone psychology manifesto. *Perspectives in Psychological Science*, 7, 221–237.

Mortenson, C. R., & Cialdini, R. B. (2010). Full-cycle social psychology for theory and application. Social and Personality Psychology Compass, 4, 53–63.

Pennebaker, J. W., Mehl, M. R., Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*, 547–577.

Ramírez-Esparza, N., Mehl, M. R., Álvarez Bermúdez, J., & Pennebaker, J. W. (2009). Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. *Journal of Research in Personality, 43*, 1–7.

Rathje, W., & Murphy, C. (2001). *Rubbish! The archaeology of garbage*. New York, NY: Harper Collins.

Reis, H. T., & Gosling, S. D. (2010). Social psychological methods outside the laboratory. In S. T. Fiske, D. T. Gilbert, & G. Lindzey, (Eds.), *Handbook of social psychology* (5th ed., Vol. 1, pp. 82–114). New York, NY: Wiley.

Sapolsky, R. (2004). *Why zebras don't get ulcers: A guide to stress, stress-related diseases and coping*. New York, NY: Henry Holt and Co.

Schlotz, W. (2012). Ambulatory psychoneuroendocrinology: Assessing salivary cortisol and other hormones in daily life. In M.R. Mehl & T.S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 193–209). New York, NY: Guilford Press.

Smyth, J., Ockenfels, M. C., Porter, L., Kirschbaum, C., Hellhammer, D. H., & Stone, A. A. (1998). Stressors and mood measured on a momentary basis are associated with salivary cortisol secretion. *Psychoneuroendocrinology, 23*, 353–370.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*, 199–202.

Stone, A. A., Reed, B. R., Neale, J. M. (1987). Changes in daily event frequency precede episodes of physical symptoms. *Journal of Human Stress, 13*, 70–74.

Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrest, L., & Grove, J. B. (1981). *Nonreactive measures in the social sciences*. Boston, MA: Houghton Mifflin Co.

White, W. B., Schulman, P., McCabe, E. J., & Dey, H. M. (1989). Average daily blood pressure, not office blood pressure, determines cardiac function in patients with hypertension. *Journal of the American Medical Association, 261*, 873–877.

Whyte, W. H. (1980). *The social life of small urban spaces*. Washington, DC: The Conservation Foundation.

Wilhelm, F.H., & Grossman, P. (2010). Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. *Biological Psychology, 84*, 552–569.

Wilhelm, P., Perrez, M., & Pawlik, K. (2012). Conducting research in daily life: A historical review. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life*. New York, NY: Guilford Press.

Wilson, R., & Gosling, S. D., & Graham, L. (2012). A review of Facebook research in the social sciences. *Perspectives on Psychological Science, 7*, 203–220.

Yarkoni, T. (2012). Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science, 21*, 391–397.

3
Psychophysiological Methods in Neuroscience

Zachary Infantolino & Gregory A. Miller

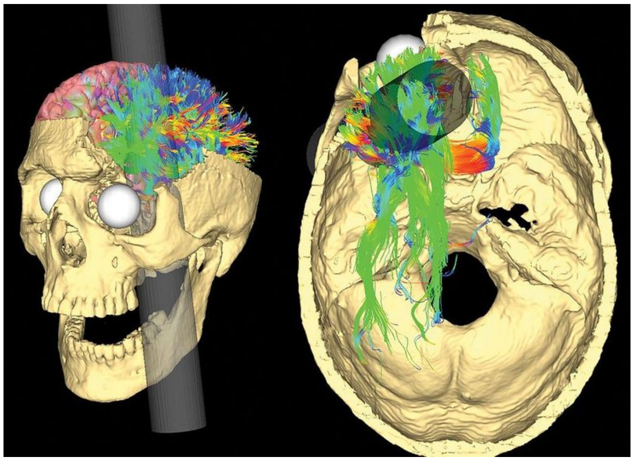
As a generally noninvasive subset of neuroscience methods, psychophysiological methods are used across a variety of disciplines in order to answer diverse questions about psychology, both mental events and behavior. Many different techniques are classified as psychophysiological. Each technique has its strengths and weaknesses, and knowing them allows researchers to decide what each offers for a particular question. Additionally, this knowledge allows research consumers to evaluate the meaning of the results in a particular experiment.

Learning Objectives

- Learn what qualifies as psychophysiology within the broader field of neuroscience.
- Review and compare several examples of psychophysiological methods.
- Understand advantages and disadvantages of different psychophysiological methods.

History

In the mid-19th century, a railroad worker named Phineas Gage was in charge of setting explosive charges for blasting through rock in order to prepare a path for railroad tracks. He would lay the charge in a hole drilled into the rock, place a fuse and sand on top of the charge, and pack it all down using a tamping iron (a solid iron rod approximately one yard long and a little over an inch in diameter). On a September afternoon when Gage was performing this task, his tamping iron caused a spark that set off the explosive prematurely, sending the tamping iron flying through the air.



Unlike other areas of the body, damage to the brain isn't localized to that specific region; injuries have widespread consequences for other areas, too. [Image: Van Horn JD, Irimia A, Torgerson CM, Chambers MC, Kikinis R, et al., https://goo.gl/wdhM4o, CC BY 2.5, https://goo.gl/0QtWcf]

Unfortunately for Gage, his head was above the hole and the tamping iron entered the side of his face, passed behind his left eye, and exited out of the top of his head, eventually landing 80 feet away. Gage lost a portion of his left frontal lobe in the accident, but survived and lived for another 12 years. What is most interesting from a psychological perspective is that Gage's personality changed as a result of this accident. He became more impulsive, he had trouble carrying out plans, and, at times, he engaged in vulgar profanity, which was out of character. This case study leads one to believe that there are specific areas of the brain that are associated with certain psychological phenomena. When studying psychology, the brain is indeed an interesting source of information. Although it would be impossible to replicate the type of damage done to Gage in the name of research, methods have developed over the years that are able to safely measure different aspects of nervous system activity in order to help researchers better understand psychology as well as the relationship between psychology and biology.

Introduction

Psychophysiology is defined as any research in which the dependent variable (what the researcher measures) is a physiological measure, and the independent variable (what the researcher manipulates) is behavioral or mental. In most cases the work is done noninvasively with awake human participants. Physiological measures take many forms and range from blood flow or neural activity in the brain to heart rate variability and eye movements. These measures can provide information about processes including emotion, cognition, and the interactions between them. In these ways, physiological measures offer a very flexible set of tools for researchers to answer questions about behavior, cognition, and health.

Psychophysiological methods are a subset of the very large domain of neuroscience methods. Many neuroscience methods are invasive, such as involving lesions of neural tissue, injection of neutrally active chemicals, or manipulation of neural activity via electrical stimulation. The present survey emphasizes noninvasive methods widely used with human subjects.

Crucially, in examining the relationship between physiology and overt behavior or mental events, psychophysiology does not attempt to replace the latter with the former. As an example, happiness is a state of pleasurable contentment and is associated with various physiological measures, but one would not say that those physiological measures *are* happiness. We can make inferences about someone's cognitive or emotional state based on his or her self-report, physiology, or overt behavior. Sometimes our interest is primarily in inferences about internal events and sometimes primarily in the physiology itself. Psychophysiology addresses both kinds of goals.

Central Nervous System (CNS)

The adult human brain only makes up about 2% (i.e. ≈ 3 pounds) of the average adult's weight, but it uses 20% of the body's energy! [Image: _DJ_, https://goo.gl/eHPH5L, CC BY-SA 2.0, https://goo.gl/rxiUsF]

This module provides an overview of several popular psychophysiological methods, though it is far from exhaustive. Each method can draw from a broad range of data-analysis strategies to provide an even more expansive set of tools. The psychophysiological methods discussed below focus on the central nervous system. Structural magnetic resonance imaging (sMRI) is a noninvasive technique that allows researchers and clinicians to view anatomical structures within a human. The participant is placed in a magnetic field that may be 66,000 times greater than the Earth's magnetic field, which causes a small portion of the atoms in his or her body to line up in the same direction. The body is then pulsed with low-energy radio frequencies that are absorbed by the atoms in the body, causing them to tip over. As these atoms return to their aligned state, they give off energy in the form of harmless electromagnetic radiation, which is measured by the machine. The machine then transforms the measured energy into a three-dimensional picture of the tissue within the body. In psychophysiology research, this image may be used to compare the size of structures in different groups of people (e.g., are areas associated with pleasure smaller in individuals with depression?) or to increase the accuracy of spatial locations as measured with functional magnetic resonance imaging (fMRI).

Functional magnetic resonance imaging (fMRI) is a method that is used to assess changes in activity of tissue, such as measuring changes in neural activity in different areas of the brain during thought. This technique builds on the principles of sMRI and also uses the property that, when neurons fire, they use energy, which must be replenished. Glucose and oxygen, two key components for energy production, are supplied to the brain from the blood stream as needed. Oxygen is transported through the blood using **hemoglobin**, which contains binding sites for oxygen. When these sites are saturated with oxygen, it is referred to as **oxygenated hemoglobin**. When the oxygen molecules have all been released from a hemoglobin molecule, it is known as **deoxygenated hemoglobin**. As a set of neurons begin firing, oxygen in the blood surrounding those neurons is consumed, leading to a reduction in oxygenated hemoglobin. The body then compensates and provides an abundance of oxygenated hemoglobin in the blood surrounding that activated neural tissue. When activity in that neural tissue declines, the level of oxygenated hemoglobin slowly returns to its original level, which typically takes several seconds.

fMRI measures the change in the concentration of oxygenated hemoglobin, which is known as the **blood-oxygen-level-dependent (BOLD)** signal. This leads to two important facts about fMRI. First, fMRI measures blood volume and blood flow, and from this we infer neural activity; fMRI does not measure neural activity directly. Second, fMRI data typically have poor **temporal resolution** (the precision of measurement with respect to time); however, when combined with sMRI, fMRI provides excellent **spatial resolution** (the ability to distinguish one object from another in space). Temporal resolution for fMRI is typically on the order of seconds, whereas its spatial resolution is on the order of millimeters. Under most conditions there is an inverse relationship between temporal and spatial resolution—one can increase temporal resolution at the expense of spatial resolution and vice versa.

This method is valuable for identifying specific areas of the brain that are associated with different physical or psychological tasks. Clinically, fMRI may be used prior to neurosurgery in order to identify areas that are associated with language so that the surgeon can avoid those areas during the operation. fMRI allows researchers to identify differential or convergent patterns of activation associated with tasks. For example, if participants are shown words on a screen and are expected to indicate the color of the letters, are the same brain areas recruited for this task if the words have emotional content or not? Does this relationship change in psychological disorders such as anxiety or depression? Is there a different pattern of activation even in the absence of overt performance differences? fMRI is an excellent tool for comparing brain activation in different tasks and/or populations. Figure 1 provides an example of results from fMRI analyses overlaid on an sMRI image. The blue and orange shapes represent areas with significant changes in the BOLD signal, thus changes in neural activation.

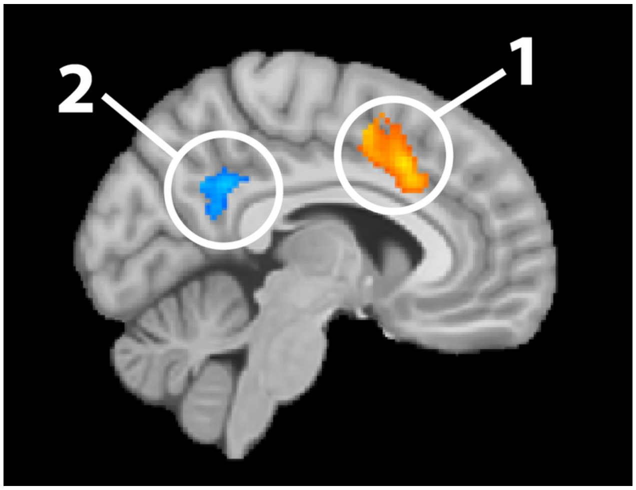


Figure 1. Example of fMRI analyses overlaid on an sMRI image. Area 1 (orange) indicates an increase in the BOLD signal, and Area 2 (blue) indicates a decrease in the BOLD signal. We infer that neural activity increased in the Area 1 and decreased in Area 2.

Electroencephalography (EEG) is another technique for studying brain activation. This technique uses at least two and sometimes up to 256 electrodes to measure the difference in electrical charge (the voltage) between pairs of points on the head. These electrodes are typically fastened to a flexible cap (similar to a swimming cap) that is placed on the participant's head. From the scalp, the electrodes measure the electrical activity that is naturally occurring within the brain. They do not introduce any new electrical activity. In contrast to fMRI, EEG measures neural activity directly, rather than a correlate of that activity.

Electrodes used in EEG can also be placed within the skull, resting directly on the brain itself. This application, called electrocorticography (ECoG), is typically used prior to medical procedures for localizing activity, such as the origin of epileptic seizures. This invasive procedure allows for more precise localization of neural activity, which is essential in medical applications. However, it is generally not justifiable to open a person's skull solely for research purposes, and instead electrodes are placed on the participant's scalp, resulting in a noninvasive technique for measuring neural activity.

Given that this electrical activity must travel through the skull and scalp before reaching the electrodes, localization of activity is less precise when measuring from the scalp, but it can still be within several millimeters when localizing activity that is near the scalp. One major advantage of EEG is its temporal resolution. Data can be recorded thousands of times per second, allowing researchers to document events that happen in less than a millisecond. EEG analyses typically investigate the change in amplitude or frequency components of the recorded EEG on an ongoing basis or averaged over dozens of trials (see Figure 2).

Magnetoencephalography (MEG) is another technique for noninvasively measuring neural activity. The flow of electrical charge (the current) associated with neural activity produces very weak magnetic fields that can be detected by sensors placed near the participant's scalp.

The number of sensors used varies from a few to several hundred. Due to the fact that the magnetic fields of interest are so small, special rooms that are shielded from magnetic fields in the environment are needed in order to avoid contamination of the signal being measured. MEG has the same excellent temporal resolution as EEG. Additionally, MEG is not as susceptible to distortions from the skull and scalp. Magnetic fields are able to pass through the hard and soft tissue relatively unchanged, thus providing better spatial resolution than EEG. MEG analytic strategies are nearly identical to those used in EEG. However, the MEG recording apparatus is much more expensive than EEG, so MEG is much less widely available.

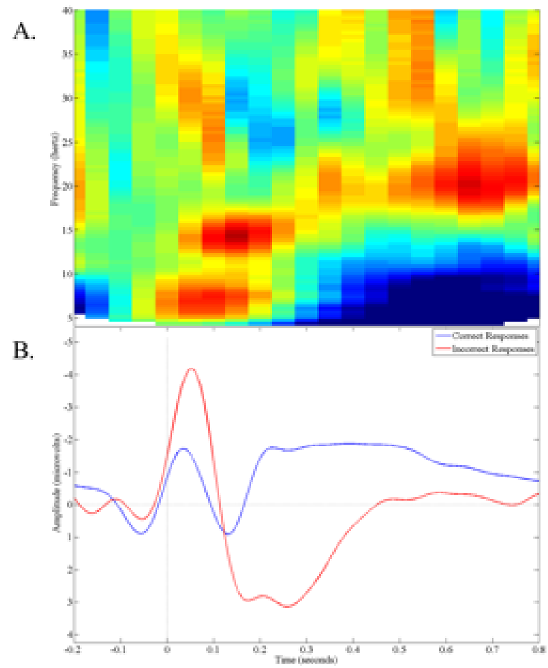


Figure 2. Example of EEG analysis output. Panel A represents changes in the relative strength of different frequencies in the EEG data over time. Panel B represents changes in the amplitude in the instantaneous EEG voltage over time.

EEG and MEG are both excellent for elucidating the temporal dynamics of neural processes. For example, if someone is reading a sentence that ends with an unexpected word (e.g., Michelle is going outside to water the book), how long after he or she reads the unexpected word does he or she recognize this as unexpected? In addition to these types of questions, EEG and MEG methods allow researchers to investigate the degree to which different parts of the brain “talk” to each other. This allows for a better understanding of brain networks, such as their role in different tasks and how they may function abnormally in psychopathology.

Positron emission tomography (PET) is a medical imaging technique that is used to measure processes in the body, including the brain. This method relies on a positron-emitting tracer atom that is introduced into the blood stream in a biologically active molecule, such as glucose, water, or ammonia. A **positron** is a particle much like an electron but with a positive charge. One example of a biologically active molecule is fludeoxyglucose, which acts similarly to glucose in the body. Fludeoxyglucose will concentrate in areas where glucose is needed—commonly areas with higher metabolic needs. Over time, this tracer molecule emits positrons, which are detected by a sensor. The spatial location of the tracer molecule in the brain can be determined based on the emitted positrons. This allows researchers to construct a three-dimensional image of the areas of the brain that have the highest metabolic needs, typically those that are most active. Images resulting from PET usually represent neural activity that has occurred over tens of minutes, which is very poor temporal resolution for some purposes. PET images are often combined with computed tomography (CT) images to improve spatial resolution, as fine as several millimeters. Tracers can also be incorporated into molecules that bind to neurotransmitter receptors, which allow researchers to answer some unique questions about the action of neurotransmitters. Unfortunately, very few research centers have the equipment required to obtain the images or the special equipment needed to create the positron-emitting tracer molecules, which typically need to be produced on site.

Transcranial magnetic stimulation (TMS) is a noninvasive method that causes **depolarization** or **hyperpolarization** in neurons near the scalp. This method is not considered psychophysiological because the independent variable is physiological, rather than the dependent. However, it does qualify as a neuroscience method because it deals with the function of the nervous system, and it can readily be combined with conventional psychophysiological methods. In TMS, a coil of wire is placed just above the participant's scalp. When electricity flows through the coil, it produces a magnetic field. This magnetic field travels through the skull and scalp and affects neurons near the surface of the brain. When the magnetic field is rapidly turned on and off, a current is induced in the neurons, leading to depolarization or hyperpolarization, depending on the number of magnetic field pulses. Single- or paired-pulse TMS depolarizes site-specific neurons in the cortex, causing them to fire. If this method is used over primary motor cortex, it can produce or block muscle activity, such as inducing a finger twitch or preventing someone from pressing a button. If used over primary visual cortex, it can produce sensations of flashes of light or impair visual processes. This has proved to be a valuable tool in studying the function and timing of specific processes such as the recognition of visual stimuli. Repetitive TMS produces effects that last longer than the initial stimulation. Depending on the intensity, coil orientation, and frequency, neural activity in the stimulated area may be either attenuated or amplified. Used in this manner, TMS is able to explore **neural plasticity**, which is the ability of connections between neurons to change. This has implications for treating psychological disorders as well as understanding long-term changes in neuronal excitability.

Peripheral Nervous System

The psychophysiological methods discussed above focus on the central nervous system. Considerable research has also focused on the **peripheral nervous system**. These methods include skin conductance, cardiovascular responses, muscle activity, pupil diameter, eye blinks, and eye movements. Skin conductance, for example, measures the electrical conductance (the inverse of resistance) between two points on the skin, which varies with the level of moisture. Sweat glands are responsible for this moisture and are controlled by the sympathetic nervous system (SNS). Increases in skin conductance can be associated with changes in psychological activity. For example, studying skin conductance allows a researcher to investigate whether psychopaths react to fearful pictures in a normal way. Skin conductance provides relatively poor temporal resolution, with the entire response typically taking several seconds to emerge and resolve. However, it is an easy way to measure SNS response to a variety of stimuli.

Cardiovascular measures include heart rate, heart rate variability, and blood pressure. The heart is innervated by the parasympathetic nervous system (PNS) and SNS. Input from the PNS decreases heart rate and contractile strength, whereas input from the SNS increases heart rate and contractile strength. Heart rate can easily be monitored using a minimum of two electrodes and is measured by counting the number of heartbeats in a given time period, such as one minute, or by assessing the time between successive heartbeats. Psychological activity can prompt increases and decreases in heart rate, often in less than a second, making heart rate a sensitive measure of cognition. Measures of heart rate variability are concerned with consistency in the time interval between heartbeats. Changes in heart rate variability are associated with stress as well as psychiatric conditions. Figure 3 is an example of an electrocardiogram, which is used to measure heart rate and heart rate variability. These cardiovascular measures allow researchers to monitor SNS and PNS reactivity to various stimuli or situations. For example, when an arachnophobe views pictures of spiders, does their heart rate increase more than that of a person not afraid of spiders?

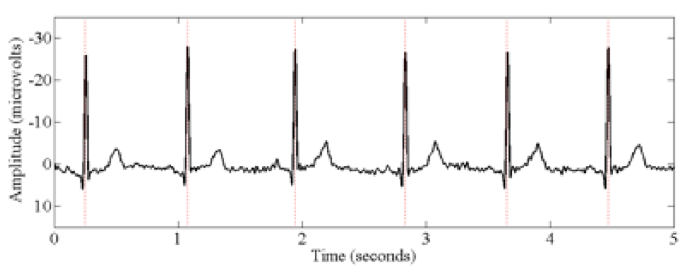


Figure 3. Example of electrocardiogram. The number of strong negative spikes in the output during a given period of time represents the heart rate, whereas the difference in the spacing between those strong negative spikes represents the heart rate variability.

Electromyography (EMG) measures electrical activity produced by skeletal muscles. Similar to EEG, EMG measures the voltage between two points. This technique can be used to determine when a participant first initiates muscle activity to engage in a motor response to a stimulus or the degree to which a participant begins to engage in an incorrect response (such as pressing the wrong button), even if it is never visibly executed. It has also been used in emotion research to identify activity in muscles that are used to produce smiles and frowns. Using EMG, it is possible to detect very small facial movements that are not observable from looking at the face. The temporal resolution of EMG is similar to that of EEG and MEG.

Valuable information can also be gleaned from eye blinks, eye movements, and pupil diameter. Eye blinks are most often assessed using EMG electrodes placed just below the eyelid, but electrical activity associated directly with eye blinks or eye movements can be measured with electrodes placed on the face near the eyes, because there is voltage across the entire eyeball. Another option for the measurement of eye movement is a camera used to record video of an eye. This video method is particularly valuable when determination of absolute direction of gaze (not just change in direction of gaze) is of interest, such as when the eyes scan a picture. With the help of a calibration period in which a participant looks at multiple, known targets, eye position is then extracted from each video frame during the main task and compared with data from the calibration phase, allowing researchers to identify the sequence, direction, and duration of gaze fixations. For example, when viewing pleasant or unpleasant images, people spend different amounts of time looking at the most arousing parts. This, in turn, can vary as a function of psychopathology. Additionally, the diameter of a participant’s pupil can be measured and recorded over time from the video record. As with heart rate, pupil diameter is controlled by competing inputs from the SNS and PNS. Pupil diameter is commonly used as an index of mental effort when performing a task.

When to Use What

As the reader, you may be wondering, how do I know what tool is right for a given question? Generally, there are no definitive answers. If you wanted to know the temperature in the morning, would you check your phone? Look outside to see how warm it looks? Ask your roommate what he or she is wearing today? Look to see what other people are wearing? There is not a single way to answer the question. The same is true for research questions. However, there are some guidelines that one can consider. For example, if you are interested in what brain structures are associated with cognitive control, you wouldn’t use peripheral nervous system measures. A technique such as fMRI or PET might be more appropriate. If you are interested in how cognitive control unfolds over time, EEG or MEG would be a good choice. If you are interested in studying the bodily response to fear in different groups of people, peripheral nervous system measures might be most appropriate. The key to deciding what method is most appropriate is properly defining the question that you are trying to answer. What aspects are most interesting? Do you care about identifying the most relevant brain structures? Temporal dynamics? Bodily responses? Then, it is important to think about the strengths and weaknesses of the different psychophysiological measures and pick one, or several, whose attributes work best for the question at hand. In fact, it is common to record several at once.

Conclusion

The outline of psychophysiological methods above provides a glimpse into the exciting techniques that are available to researchers studying a broad range of topics from clinical to social to cognitive psychology. Some of the most interesting psychophysiological studies use several methods, such as in sleep assessments or multimodal neuroimaging. Psychophysiological methods have applications outside of mainstream psychology in areas where psychological phenomena are central, such as economics, health-related decision making, and brain–computer interfaces. Examples of applications for each method are provided above, but this list is by no means exhaustive. Furthermore, the field is continually evolving, with new methods and new applications being developed. The wide variety of methods and applications provide virtually limitless possibilities for researchers.

Outside Resources

Book: Luck, S. J. (2005). An introduction to the event-related potential technique. Cambridge, MA: MIT Press.

Book: Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). Handbook of functional MRI data analysis. New York: Cambridge University Press.

Web: For visualizations on MRI physics (requires a free registration):
http://www.imaios.com/en/e-Courses/e-MRI/NMR/

Discussion Questions

1. Pick a psychological phenomenon that you would like to know more about. What specific hypothesis would you like to test? What psychophysiological methods might be appropriate for testing this hypothesis and why?
2. What types of questions would require high spatial resolution in measuring brain activity? What types of questions would require high temporal resolution?
3. Take the hypothesis you picked in the first question, and choose what you think would be the best psychophysiological method. What additional information could you obtain using a complementary method? For example, if you want to learn about memory, what two methods could you use that would each provide you with distinct information?
4. The popular press has shown an increasing interest in findings that contain images of brains and neuroscience language. Studies have shown that people often find presentations of results that contain these features more convincing than presentations of results that do not, even if the actual results are the same. Why would images of the brain and neuroscience language be more convincing to people? Given that results with these features are more convincing, what do you think is the researcher's responsibility in reporting results with brain images and neuroscience language?
5. Many claims in the popular press attempt to reduce complex psychological phenomena to biological events. For example, you may have heard it said that schizophrenia is a brain disorder or that depression is simply a chemical imbalance. However, this type of "reductionism" so far does not appear to be tenable. There has been surprisingly little discussion of possible causal relationships, in either direction, between biological and psychological phenomena. We are aware of no such documented causal mechanisms. Do

you think that it will ever be possible to explain how a change in biology can result in a change of a psychological phenomenon, or vice versa?

Vocabulary

Blood-oxygen-level-dependent (BOLD)
The signal typically measured in fMRI that results from changes in the ratio of oxygenated hemoglobin to deoxygenated hemoglobin in the blood.

Central nervous system
The part of the nervous system that consists of the brain and spinal cord.

Deoxygenated hemoglobin
Hemoglobin not carrying oxygen.

Depolarization
A change in a cell's membrane potential, making the inside of the cell more positive and increasing the chance of an action potential.

Hemoglobin
The oxygen-carrying portion of a red blood cell.

Hyperpolarization
A change in a cell's membrane potential, making the inside of the cell more negative and decreasing the chance of an action potential.

Invasive Procedure
A procedure that involves the skin being broken or an instrument or chemical being introduced into a body cavity.

Lesions
Abnormalities in the tissue of an organism usually caused by disease or trauma.

Neural plasticity
The ability of synapses and neural pathways to change over time and adapt to changes in neural process, behavior, or environment.

Neuroscience methods
A research method that deals with the structure or function of the nervous system and brain.

Noninvasive procedure

A procedure that does not require the insertion of an instrument or chemical through the skin or into a body cavity.

Oxygenated hemoglobin
Hemoglobin carrying oxygen.

Parasympathetic nervous system (PNS)
One of the two major divisions of the autonomic nervous system, responsible for stimulation of “rest and digest” activities.

Peripheral nervous system
The part of the nervous system that is outside the brain and spinal cord.

Positron
A particle having the same mass and numerically equal but positive charge as an electron.

Psychophysiological methods
Any research method in which the dependent variable is a physiological measure and the independent variable is behavioral or mental (such as memory).

Spatial resolution
The degree to which one can separate a single object in space from another.

Sympathetic nervous system (SNS)
One of the two major divisions of the autonomic nervous system, responsible for stimulation of “fight or flight” activities.

Temporal resolution
The degree to which one can separate a single point in time from another.

Voltage
The difference in electric charge between two points.

4
History of Psychology

David B. Baker & Heather Sperry

This module provides an introduction and overview of the historical development of the science and practice of psychology in America. Ever-increasing specialization within the field often makes it difficult to discern the common roots from which the field of psychology has evolved. By exploring this shared past, students will be better able to understand how psychology has developed into the discipline we know today.

Learning Objectives

- Describe the precursors to the establishment of the science of psychology.
- Identify key individuals and events in the history of American psychology.
- Describe the rise of professional psychology in America.
- Develop a basic understanding of the processes of scientific development and change.
- Recognize the role of women and people of color in the history of American psychology.

Introduction

It is always a difficult question to ask, where to begin to tell the story of the history of psychology. Some would start with ancient Greece; others would look to a demarcation in the late 19th century when the science of psychology was formally proposed and instituted. These two perspectives, and all that is in between, are appropriate for describing a history of psychology. The interested student will have no trouble finding an abundance of resources on all of these time frames and perspectives (Goodwin, 2011; Leahey, 2012; Schultz & Schultz, 2007). For the purposes of this module, we will examine the development of psychology in America and use the mid-19th century as our starting point. For the sake of convenience, we refer to this as a history of modern psychology.

The earliest records of a psychological experiment go all the way back to the Pharaoh Psamtik I of Egypt in the 7th Century B.C. [Image: Neithsabes, CC0 Public Domain, https://goo.gl/m25gce]

Psychology is an exciting field and the history of psychology offers the opportunity to make sense of how it has grown and developed. The history of psychology also provides perspective. Rather than a dry collection of names and dates, the history of psychology tells us about the important intersection of time and place that defines who we are. Consider what happens when you meet someone for the first time. The conversation usually begins with a series of questions such as, “Where did you grow up?” “How long have you lived here?” “Where did you go to school?” The importance of history in defining who we are cannot be overstated. Whether you are seeing a physician, talking with a counselor, or applying for a job, everything begins with a history. The same is true for studying the history of psychology; getting a history of the field helps to make sense of where we are and how we got here.

A Prehistory of Psychology

Precursors to American psychology can be found in philosophy and physiology. Philosophers such as John Locke (1632–1704) and Thomas Reid (1710–1796) promoted **empiricism**, the idea that all knowledge comes from experience. The work of Locke, Reid, and others emphasized the role of the human observer and the primacy of the senses in defining how the mind comes to acquire knowledge. In American colleges and universities in the early 1800s, these principles were taught as courses on mental and moral philosophy. Most often these courses taught about the mind based on the faculties of intellect, will, and the senses (Fuchs, 2000).

Physiology and Psychophysics

Philosophical questions about the nature of mind and knowledge were matched in the 19th century by physiological investigations of the sensory systems of the human observer. German physiologist Hermann von Helmholtz (1821–1894) measured the speed of the **neural impulse** and explored the physiology of hearing and vision. His work indicated that our senses can deceive us and are not a mirror of the external world. Such work showed that even though the human senses were fallible, the mind could be measured using the methods of science. In all, it suggested that a science of psychology was feasible.

An important implication of Helmholtz's work was that there is a psychological reality and a physical reality and that the two are not identical. This was not a new idea; philosophers like John Locke had written extensively on the topic, and in the 19th century, philosophical speculation about the nature of mind became subject to the rigors of science.

The question of the relationship between the mental (experiences of the senses) and the material (external reality) was investigated by a number of German researchers including Ernst Weber and Gustav Fechner. Their work was called **psychophysics**, and it introduced methods for measuring the relationship between physical stimuli and human perception that would serve as the basis for the new science of psychology (Fancher & Rutherford, 2011).

The formal development of modern psychology is usually credited to the work of German physician, physiologist, and philosopher Wilhelm Wundt (1832–1920). Wundt helped to establish the field of experimental psychology by serving as a strong promoter of the idea that psychology could be an experimental field and by providing classes, textbooks, and a laboratory for training students. In 1875, he joined the faculty at the University of Leipzig and quickly began to make plans for the creation of a program of experimental psychology. In 1879, he complemented his lectures on experimental psychology with a laboratory experience: an event that has served as the popular date for the establishment of the science of psychology.

Wilhelm Wundt is considered one of the founding figures of modern psychology. [CC0 Public Domain, https://goo.gl/m25gce]

The response to the new science was immediate and global. Wundt attracted students from around the world to study the new experimental psychology and work in his lab. Students were trained to offer detailed self-reports of their reactions to various stimuli, a procedure known as **introspection**. The goal was to identify the elements of **consciousness**. In addition to the study of sensation and perception, research was done on mental chronometry, more commonly known as reaction time. The work of Wundt and his students demonstrated that the mind could be measured and the nature of consciousness could be revealed through scientific means. It was an exciting proposition, and one that found great interest in America. After the opening of Wundt's lab in 1879, it took just four years for the first psychology laboratory to open in the United States (Benjamin, 2007).

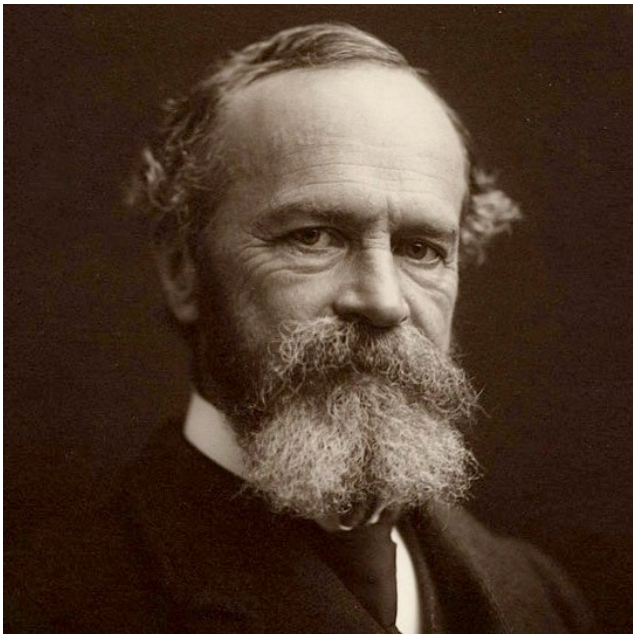
Scientific Psychology Comes to the United States

Wundt's version of psychology arrived in America most visibly through the work of Edward Bradford Titchener (1867–1927). A student of Wundt's, Titchener brought to America a brand of experimental psychology referred to as "**structuralism**." Structuralists were interested in the contents of the mind—what the mind is. For Titchener, the general adult mind was the proper focus for the new psychology, and he excluded from study those with mental deficiencies, children, and animals (Evans, 1972; Titchener, 1909).

Experimental psychology spread rather rapidly throughout North America. By 1900, there were more than 40 laboratories in the United States and Canada (Benjamin, 2000). Psychology in America also organized early with the establishment of the American Psychological Association (APA) in 1892. Titchener felt that this new organization did not adequately represent the interests of experimental psychology, so, in 1904, he organized a group of colleagues to create what is now known as the Society of Experimental Psychologists (Goodwin, 1985). The group met annually to discuss research in experimental psychology. Reflecting the times, women researchers were not invited (or welcome). It is interesting to note that Titchener's first doctoral student was a woman, Margaret Floy Washburn (1871–1939). Despite many barriers, in 1894, Washburn became the first woman in America to earn a Ph.D. in psychology and, in 1921, only the second woman to be elected president of the American Psychological Association (Scarborough & Furumoto, 1987).

Striking a balance between the science and practice of psychology continues to this day. In 1988, the American Psychological Society (now known as the Association for Psychological Science) was founded with the central mission of advancing psychological science.

Toward a Functional Psychology



William James was one of the leading figures in a new perspective on psychology called functionalism. [Image: Notman Studios, CC0 Public Domain, https://goo.gl/m25gce]

While Titchener and his followers adhered to a structural psychology, others in America were pursuing different approaches. William James, G. Stanley Hall, and James McKeen Cattell were among a group that became identified with "**functionalism**." Influenced by Darwin's evolutionary theory, functionalists were interested in the activities of the mind—what the mind does. An interest in functionalism opened the way for the study of a wide range of approaches, including animal and comparative psychology (Benjamin, 2007).

William James (1842–1910) is regarded as writing perhaps the most influential and important book in the field of psychology, *Principles of Psychology,* published in 1890. Opposed to the reductionist ideas of Titchener, James proposed that consciousness is ongoing and continuous; it cannot be isolated and reduced to elements. For James, consciousness helped us adapt to our environment in such ways as allowing us to make choices and have personal responsibility over those choices.

At Harvard, James occupied a position of authority and respect in psychology and philosophy. Through his teaching and writing, he influenced psychology for generations. One of his students, Mary Whiton Calkins (1863–1930), faced many of the challenges that confronted Margaret Floy Washburn and other women interested in pursuing graduate education in psychology. With much persistence, Calkins was able to study with James at Harvard. She eventually completed all the requirements for the doctoral degree, but Harvard refused to grant her a diploma because she was a woman. Despite these challenges, Calkins went on to become an accomplished researcher and the first woman elected president of the American Psychological Association in 1905 (Scarborough & Furumoto, 1987).

G. Stanley Hall (1844–1924) made substantial and lasting contributions to the establishment of psychology in the United States. At Johns Hopkins University, he founded the first psychological laboratory in America in 1883. In 1887, he created the first journal of psychology in America, *American Journal of Psychology*. In 1892, he founded the American Psychological Association (APA); in 1909, he invited and hosted Freud at Clark University (the only time Freud visited America). Influenced by evolutionary theory, Hall was interested in the process of adaptation and human development. Using surveys and questionnaires to study children, Hall wrote extensively on child development and education. While graduate education in psychology was restricted for women in Hall's time, it was all but non-existent for African Americans. In another first, Hall mentored Francis Cecil Sumner (1895–1954) who, in 1920, became the first African American to earn a Ph.D. in psychology in America (Guthrie, 2003).

James McKeen Cattell (1860–1944) received his Ph.D. with Wundt but quickly turned his interests to the assessment of **individual differences**. Influenced by the work of Darwin's cousin, Frances Galton, Cattell believed that mental abilities such as intelligence were inherited and could be measured using mental tests. Like Galton, he believed society was better served by identifying those with superior intelligence and supported efforts to encourage them to reproduce. Such beliefs were associated with **eugenics** (the promotion of selective breeding) and fueled early debates about the contributions of heredity and environment in defining who we are. At Columbia University, Cattell developed a department of psychology that became world famous also promoting psychological science through advocacy and as a publisher of scientific journals and reference works (Fancher, 1987; Sokal, 1980).

The Growth of Psychology

Throughout the first half of the 20th century, psychology continued to grow and flourish in America. It was large enough to accommodate varying points of view on the nature of mind and behavior. **Gestalt psychology** is a good example. The Gestalt movement began in Germany with the work of Max Wertheimer (1880–1943). Opposed to the reductionist approach of Wundt's laboratory psychology, Wertheimer and his colleagues Kurt Koffka (1886–1941), Wolfgang Kohler (1887–1967), and Kurt Lewin (1890–1947) believed that studying the whole of any experience was richer than studying individual aspects of that experience. The saying "the whole is greater than the sum of its parts" is a Gestalt perspective. Consider that a melody is an additional element beyond the collection of notes that comprise it. The Gestalt psychologists proposed that the mind often processes information simultaneously rather than sequentially. For instance, when you look at a photograph, you see a whole image, not just a collection of pixels of color. Using Gestalt principles, Wertheimer and his colleagues also explored the nature of learning and thinking. Most of the German Gestalt psychologists were Jewish and were forced to flee the Nazi regime due to the threats posed on both academic and personal freedoms. In America, they were able to introduce a new audience to the Gestalt perspective, demonstrating how it could be applied to perception and learning (Wertheimer, 1938). In many ways, the work of the Gestalt psychologists served as a precursor to the rise of **cognitive psychology** in America (Benjamin, 2007).

Behaviorism emerged early in the 20th century and became a major force in American psychology. Championed by psychologists such as John B. Watson (1878–1958) and B. F. Skinner (1904–1990), behaviorism rejected any reference to mind and viewed overt and observable behavior as the proper subject matter of psychology. Through the scientific study of behavior, it was hoped that laws of learning could be derived that would promote the prediction and control of behavior. Russian physiologist Ivan Pavlov (1849–1936) influenced early behaviorism in America. His work on conditioned learning, popularly referred to as classical conditioning, provided support for the notion that learning and behavior were controlled by events in the environment and could be explained with no reference to mind or consciousness (Fancher, 1987).

For decades, behaviorism dominated American psychology. By the 1960s, psychologists began to recognize that behaviorism was unable to fully explain human behavior because it neglected mental processes. The turn toward a cognitive psychology was not new. In the 1930s, British psychologist Frederic C. Bartlett (1886–1969) explored the idea of the constructive mind, recognizing that people use their past experiences to construct frameworks in which to understand new experiences. Some of the major pioneers in American cognitive psychology include Jerome Bruner (1915–), Roger Brown (1925–1997), and George Miller (1920–2012). In the 1950s, Bruner conducted pioneering studies on cognitive aspects of sensation and perception. Brown conducted original research on language and memory, coined the term "**flashbulb memory**," and figured out how to study the **tip-of-the-tongue phenomenon** (Benjamin, 2007). Miller's research on working memory is legendary. His 1956 paper "The Magic Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information"is one of the most highly cited papers in psychology. A popular interpretation of Miller's research was that the number of bits of information an average human can hold in **working memory** is 7 ± 2. Around the same time, the study of computer science was growing and was used as an analogy to explore and understand how the mind works. The work of Miller and others in the 1950s and 1960s has inspired tremendous interest in cognition and neuroscience, both of which dominate much of contemporary American psychology.

Applied Psychology in America

In America, there has always been an interest in the application of psychology to everyday life. Mental testing is an important example. Modern intelligence tests were developed by the French psychologist Alfred Binet (1857–1911). His goal was to develop a test that would identify schoolchildren in need of educational support. His test, which included tasks of reasoning and problem solving, was introduced in the United States by Henry Goddard (1866–1957) and later standardized by Lewis Terman (1877–1956) at Stanford University. The assessment and meaning of intelligence has fueled debates in American psychology and society for nearly 100 years. Much of this is captured in the nature-nurture debate that raises questions about the relative contributions of heredity and environment in determining intelligence (Fancher, 1987).

Applied psychology was not limited to mental testing. What psychologists were learning in their laboratories was applied in many settings including the military, business, industry, and education. The early 20th century was witness to rapid advances in applied psychology. Hugo Munsterberg (1863–1916) of Harvard University made contributions to such areas as employee selection, eyewitness testimony, and psychotherapy. Walter D. Scott (1869–1955) and Harry Hollingworth (1880–1956) produced original work on the psychology of advertising and marketing. Lillian Gilbreth (1878–1972) was a pioneer in industrial psychology and engineering psychology. Working with her husband, Frank, they promoted the use of time and motion studies to improve efficiency in industry. Lillian also brought the efficiency movement to the home, designing kitchens and appliances including the pop-up trashcan and refrigerator door shelving. Their psychology of efficiency also found plenty of applications at home with their 12 children. The experience served as the inspiration for the movie *Cheaper by the Dozen* (Benjamin, 2007).

Clinical psychology was also an early application of experimental psychology in America. Lightner Witmer (1867–1956) received his Ph.D. in experimental psychology with Wilhelm Wundt and returned to the University of Pennsylvania, where he opened a psychological clinic in 1896. Witmer believed that because psychology dealt with the study of sensation and perception, it should be of value in treating children with learning and behavioral problems. He is credited as the founder of both clinical and school psychology (Benjamin & Baker, 2004).

Although this is what most people see in their mind's eye when asked to envision a "psychologist" the APA recognizes as many as 58 different divisions of psychology. [Image: Bliusa, https://goo.gl/yrSUCr, CC BY-SA 4.0, https://goo.gl/6pvNbx]

Psychology as a Profession

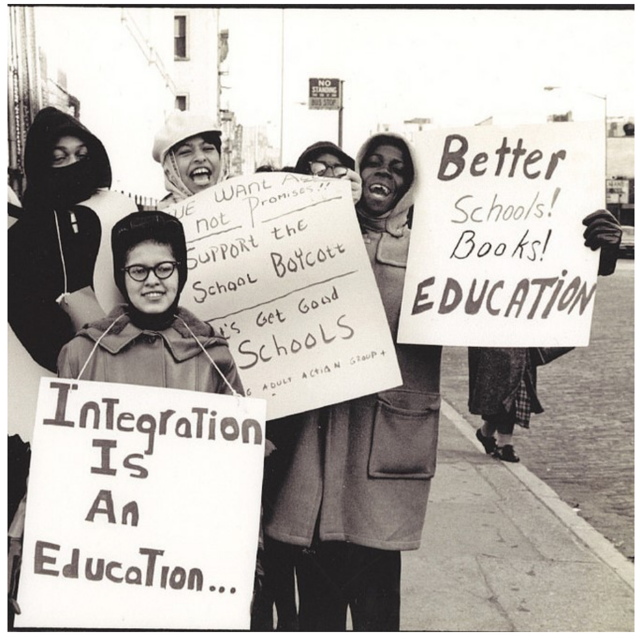
As the roles of psychologists and the needs of the public continued to change, it was necessary for psychology to begin to define itself as a profession. Without standards for training and practice, anyone could use the title psychologist and offer services to the public. As early as 1917, applied psychologists organized to create standards for education, training, and licensure. By the 1930s, these efforts led to the creation of the American Association for Applied Psychology (AAAP). While the American Psychological Association (APA) represented the interests of academic psychologists, AAAP served those in education, industry, consulting, and clinical work.

The advent of WWII changed everything. The psychiatric casualties of war were staggering, and there were simply not enough mental health professionals to meet the need. Recognizing the shortage, the federal government urged the AAAP and APA to work together to meet the mental health needs of the nation. The result was the merging of the AAAP and the APA and a focus on the training of professional psychologists. Through the provisions of National Mental Health Act of 1946, funding was made available that allowed the APA, the Veterans Administration, and the Public Health Service to work together to develop training programs that would produce clinical psychologists. These efforts led to the convening of the Boulder Conference on Graduate Education in Clinical Psychology in 1949 in Boulder, Colorado. The meeting launched doctoral training in psychology and gave us the **scientist-practitioner model** of training. Similar meetings also helped launch doctoral training programs in counseling and school psychology. Throughout the second half of the 20th century, alternatives to Boulder have been debated. In 1973, the Vail Conference on Professional Training in Psychology proposed the **scholar-practitioner model** and the Psy.D. degree (Doctor of Psychology). It is a training model that emphasizes clinical training and practice that has become more common (Cautin & Baker, in press).

Psychology and Society

Given that psychology deals with the human condition, it is not surprising that psychologists would involve themselves in social issues. For more than a century, psychology and psychologists have been agents of social action and change. Using the methods and tools of science, psychologists have challenged assumptions, stereotypes, and stigma. Founded in 1936, the Society for the Psychological Study of Social Issues (SPSSI) has supported research and action on a wide range of social issues. Individually, there have been many psychologists whose efforts have promoted social change. Helen Thompson Woolley (1874–1947) and Leta S. Hollingworth (1886–1939) were pioneers in research on the psychology of sex differences. Working in the early 20th century, when women's rights were marginalized, Thompson examined the assumption that women were overemotional compared to men and found that emotion did not influence women's decisions any more than it did men's. Hollingworth found that menstruation did not negatively impact women's cognitive or motor abilities. Such work combatted harmful stereotypes and showed that psychological research could contribute to social change (Scarborough & Furumoto, 1987).

Among the first generation of African American psychologists, Mamie Phipps Clark (1917–1983) and her husband Kenneth Clark (1914–2005) studied the psychology of race and demonstrated the ways in which school segregation negatively impacted the self-esteem of African American children. Their research was influential in the 1954 Supreme Court ruling in the case of *Brown v. Board of Education,* which ended school segregation (Guthrie, 2003). In psychology, greater advocacy for issues impacting the African American community were advanced by the creation of the Association of Black Psychologists (ABPsi) in 1968.



Mamie Phipps Clark and Kenneth Clark studied the negative impacts of segregated education on African-American children. [Image: Penn State Special Collection, https://goo.gl/WP7Dgc, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

In 1957, psychologist Evelyn Hooker (1907–1996) published the paper "The Adjustment of the Male Overt Homosexual," reporting on her research that showed no significant differences in psychological adjustment between homosexual and heterosexual men. Her research helped to de-pathologize homosexuality and contributed to the decision by the American Psychiatric Association to remove homosexuality from the Diagnostic and Statistical Manual of Mental Disorders in 1973 (Garnets & Kimmel, 2003).

Conclusion

Growth and expansion have been a constant in American psychology. In the latter part of the 20th century, areas such as social, developmental, and personality psychology made major contributions to our understanding of what it means to be human. Today neuroscience is enjoying tremendous interest and growth.

As mentioned at the beginning of the module, it is a challenge to cover all the history of psychology in such a short space. Errors of omission and commission are likely in such a selective review. The history of psychology helps to set a stage upon which the story of psychology can be told. This brief summary provides some glimpse into the depth and rich content offered by the history of psychology. The learning modules in the Noba psychology collection are all elaborations on the foundation created by our shared past. It is hoped that you will be able to see these connections and have a greater understanding and appreciation for both the unity and diversity of the field of psychology.

Timeline

1600s – Rise of empiricism emphasizing centrality of human observer in acquiring knowledge

1850s - Helmholz measures neural impulse / Psychophysics studied by Weber & Fechner

1859 - Publication of Darwin's *Origin of Species*

1879 - Wundt opens lab for experimental psychology

1883 - First psychology lab opens in the United States

1887 – First American psychology journal is published: *American Journal of Psychology*

1890 – James publishes *Principles of Psychology*

1892 – APA established

1894 – Margaret Floy Washburn is first U.S. woman to earn Ph.D. in psychology

1904 - Founding of Titchener's experimentalists

1905 - Mary Whiton Calkins is first woman president of APA

1909 – Freud’s only visit to the United States

1913 - John Watson calls for a psychology of behavior

1920 – Francis Cecil Sumner is first African American to earn Ph.D. in psychology

1921 – Margaret Floy Washburn is second woman president of APA

1930s – Creation and growth of the American Association for Applied Psychology (AAAP) / Gestalt psychology comes to America

1936- Founding of The Society for the Psychological Study of Social Issues

1940s – Behaviorism dominates American psychology

1946 – National Mental Health Act

1949 – Boulder Conference on Graduate Education in Clinical Psychology

1950s – Cognitive psychology gains popularity

1954 – *Brown v. Board of Education*

1957 – Evelyn Hooker publishes T*he Adjustment of the Male Overt Homosexual*

1968 – Founding of the Association of Black Psychologists

1973 – Psy.D. proposed at the Vail Conference on Professional Training in Psychology

1988 – Founding of the American Psychological Society (now known as the Association for Psychological Science)

Outside Resources

Podcast: History of Psychology Podcast Series
http://www.yorku.ca/christo/podcasts/

Web: Advances in the History of Psychology
http://ahp.apps01.yorku.ca/

Web: Center for the History of Psychology
http://www.uakron.edu/chp

Web: Classics in the History of Psychology
http://psychclassics.yorku.ca/

Web: Psychology's Feminist Voices
http://www.feministvoices.com/

Web: This Week in the History of Psychology
http://www.yorku.ca/christo/podcasts/

Discussion Questions

1. Why was psychophysics important to the development of psychology as a science?
2. How have psychologists participated in the advancement of social issues?
3. Name some ways in which psychology began to be applied to the general public and everyday problems.
4. Describe functionalism and structuralism and their influences on behaviorism and cognitive psychology.

Vocabulary

Behaviorism
The study of behavior.

Cognitive psychology
The study of mental processes.

Consciousness
Awareness of ourselves and our environment.

Empiricism
The belief that knowledge comes from experience.

Eugenics
The practice of selective breeding to promote desired traits.

Flashbulb memory
A highly detailed and vivid memory of an emotionally significant event.

Functionalism
A school of American psychology that focused on the utility of consciousness.

Gestalt psychology
An attempt to study the unity of experience.

Individual differences
Ways in which people differ in terms of their behavior, emotion, cognition, and development.

Introspection
A method of focusing on internal processes.

Neural impulse
An electro-chemical signal that enables neurons to communicate.

Practitioner-Scholar Model
A model of training of professional psychologists that emphasizes clinical practice.

Psychophysics
Study of the relationships between physical stimuli and the perception of those stimuli.

Realism
A point of view that emphasizes the importance of the senses in providing knowledge of the external world.

Scientist-practitioner model
A model of training of professional psychologists that emphasizes the development of both research and clinical skills.

Structuralism
A school of American psychology that sought to describe the elements of conscious experience.

Tip-of-the-tongue phenomenon
The inability to pull a word from memory even though there is the sensation that that word is available.

References

Benjamin, L. T. (2007). *A brief history of modern psychology*. Malden, MA: Blackwell Publishing.

Benjamin, L. T. (2000). The psychology laboratory at the turn of the 20th century. *American Psychologist*, 55, 318–321.

Benjamin, L. T., & Baker, D. B. (2004). *From séance to science: A history of the profession of psychology in America*. Belmont, CA: Wadsworth/Thomson Learning.

Cautin, R., & Baker, D. B. (in press). A history of education and training in professional psychology. In B. Johnson & N. Kaslow (Eds.), *Oxford handbook of education and training in professional psychology*. New York, NY: Oxford University Press.

Evans, R. B. (1972). E. B. Titchener and his lost system. Journal of the History of the Behavioral Sciences, 8, 168–180.

Fancher, R. E. (1987). *The intelligence men: Makers of the IQ controversy*. New York, NY: W.W. Norton & Company.

Fancher, R. E., & Rutherford, A. (2011). *Pioneers of psychology: A history* (4th ed.). New York, NY: W.W. Norton & Company.

Fuchs, A. H. (2000). Contributions of American mental philosophers to psychology in the United States. *History of Psychology*, 3, 3–19.

Garnets, L., & Kimmel, D. C. (2003). What a light it shed: The life of Evelyn Hooker. In L. Garnets & D. C. Kimmel (Eds.), *Psychological perspectives on gay, lesbian, and bisexual experiences* (2nd ed., pp. 31–49). New York, NY: Columbia University Press.

Goodwin, C. J. (2011). *A history of modern psychology* (4th ed.). Hoboken, NJ: Wiley.

Goodwin, C. J. (1985). On the origins of Titchener's experimentalists. *Journal of the History of the Behavioral Sciences*, 21, 383–389.

Guthrie, R. V. (2003). *Even the rat was white: A historical view of psychology* (2nd ed.). Boston, MA: Allyn & Bacon.

Leahey, T. H. (2012). *A history of psychology: From antiquity to modernity* (7th ed.). Upper Saddle River, NJ: Pearson Education.

Scarborough, E. & Furumoto, L. (1987). *The untold lives: The first generation of American women psychologists*. New York, NY: Columbia University Press.

Shultz, D. P., & Schultz, S. E. (2007). *A history of modern psychology* (9th ed.). Stanford, CT: Cengage Learning.

Sokal, M. M. (1980). Science and James McKeen Cattell. *Science*, 209, 43–52.

Titchener, E. B. (1909). *A text-book of psychology*. New York, NY: Macmillan.

Wertheimer, M. (1938). Gestalt theory. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (1-11). New York, NY: Harcourt.

Biological Basis of Behavior

5

The Brain

Diane Beck & Evelina Tapia

The human brain is responsible for all behaviors, thoughts, and experiences described in this textbook. This module provides an introductory overview of the brain, including some basic neuroanatomy, and brief descriptions of the neuroscience methods used to study it.

Learning Objectives

- Name and describe the basic function of the brain stem, cerebellum, and cerebral hemispheres.
- Name and describe the basic function of the four cerebral lobes: occipital, temporal, parietal, and frontal cortex.
- Describe a split-brain patient and at least two important aspects of brain function that these patients reveal.
- Distinguish between gray and white matter of the cerebral hemispheres.
- Name and describe the most common approaches to studying the human brain.
- Distinguish among four neuroimaging methods: PET, fMRI, EEG, and DOI.
- Describe the difference between spatial and temporal resolution with regard to brain function.

Introduction

Any textbook on psychology would be incomplete without reference to the brain. Every behavior, thought, or experience described in the other modules must be implemented in the brain. A detailed understanding of the human brain can help us make sense of human experience and behavior. For example, one well-established fact about human cognition is that it is limited. We cannot do two complex tasks at once: We cannot read and carry on a conversation at the same time, text and drive, or surf the Internet while listening to a lecture, at least not successfully or safely. We cannot even pat our head and rub our stomach at the same time (with exceptions, see "A Brain Divided"). Why is this? Many people have suggested that such limitations reflect the fact that the behaviors draw on the same resource; if one behavior uses up most of the resource there is not enough resource left for the other. But what might this limited resource be in the brain?

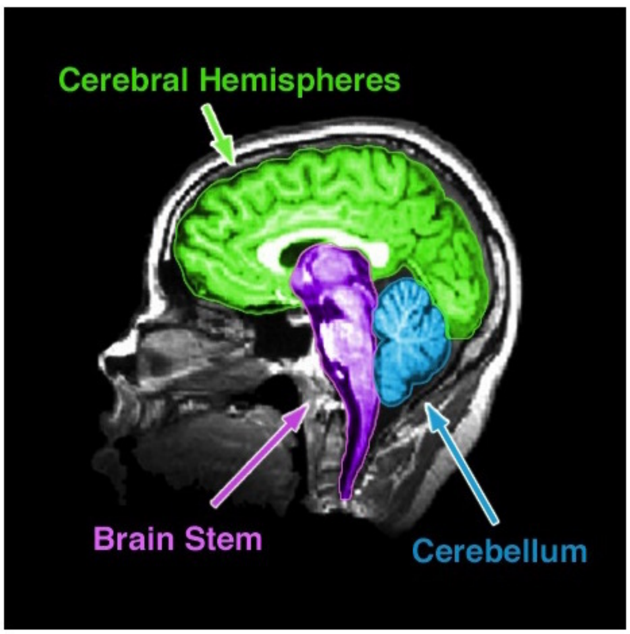


Figure 1. An MRI of the human brain delineating three major structures: the cerebral hemispheres, brain stem, and cerebellum.

The brain uses oxygen and glucose, delivered via the blood. The brain is a large consumer of these **metabolites**, using 20% of the oxygen and calories we consume despite being only 2% of our total weight. However, as long as we are not oxygen-deprived or malnourished, we have more than enough oxygen and glucose to fuel the brain. Thus, insufficient "brain fuel" cannot explain our limited capacity. Nor is it likely that our limitations reflect too few neurons. The average human brain contains 86 billion neurons. It is also not the case that we use only 10% of our brain, a myth that was likely started to imply we had untapped potential. Modern neuroimaging (see "Studying the Human Brain") has shown that we use all parts of brain, just at different times, and certainly more than 10% at any one time.

If we have an abundance of brain fuel and neurons, how can we explain our limited cognitive abilities? Why can't we do more at once? The most likely explanation is the way these neurons are wired up. We know, for instance, that many neurons in the visual cortex (the part of the brain responsible for processing visual information) are hooked up in such a way as to inhibit each other (Beck & Kastner, 2009). When one neuron fires, it suppresses the firing of other nearby neurons. If two neurons that are hooked up in an inhibitory way both fire, then neither neuron can fire as vigorously as it would otherwise. This competitive behavior among neurons limits how much visual information the brain can respond to at the same time. Similar kinds of competitive wiring among neurons may underlie many of our limitations. Thus, although talking about limited resources provides an intuitive description of our limited capacity behavior, a detailed understanding of the brain suggests that our limitations more likely reflect the complex way in which neurons talk to each other rather than the depletion of any specific resource.

The Anatomy of the Brain

There are many ways to subdivide the mammalian brain, resulting in some inconsistent and ambiguous **nomenclature** over the history of neuroanatomy (Swanson, 2000). For simplicity, we will divide the brain into three basic parts: the brain stem, cerebellum, and cerebral hemispheres (see Figure 1). In Figure 2, however, we depict other prominent groupings (Swanson, 2000) of the six major subdivisions of the brain (Kandal, Schwartz, & Jessell, 2000).

Brain Stem

The **brain stem** is sometimes referred to as the "trunk" of the brain. It is responsible for many of the neural functions that keep us alive, including regulating our respiration (breathing), heart rate, and digestion. In keeping with its function, if a patient sustains severe damage to the brain stem he or she will require "life support" (i.e., machines are used to keep him or her alive). Because of its vital role in survival, in many countries a person who has lost brain stem function is said to be "brain dead," although other countries require significant tissue loss in the cortex (of the cerebral hemispheres), which is responsible for our conscious experience, for the same diagnosis. The brain stem includes the medulla, pons, midbrain, and diencephalon (which consists of thalamus and hypothalamus). Collectively, these regions also are involved in our sleep–wake cycle, some sensory and motor function, as well as growth and other hormonal behaviors.

Cerebellum

The **cerebellum** is the distinctive structure at the back of the brain. The Greek philosopher and scientist Aristotle aptly referred to it as the "small brain" ("parencephalon" in Greek, "cerebellum" in Latin) in order to distinguish it from the "large brain" ("encephalon" in Greek, "**cerebrum**" in Latin). The cerebellum is critical for coordinated movement and posture. More recently, neuroimaging studies (see "Studying the Human Brain") have implicated it in a range of cognitive abilities, including language. It is perhaps not surprising that the cerebellum's

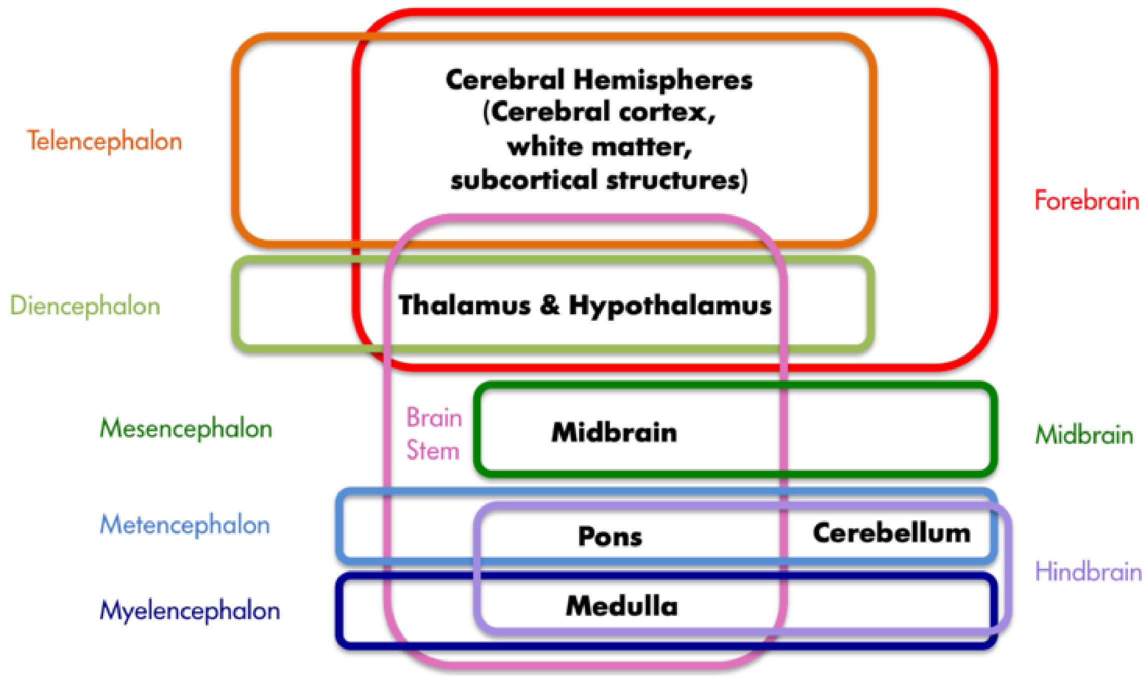


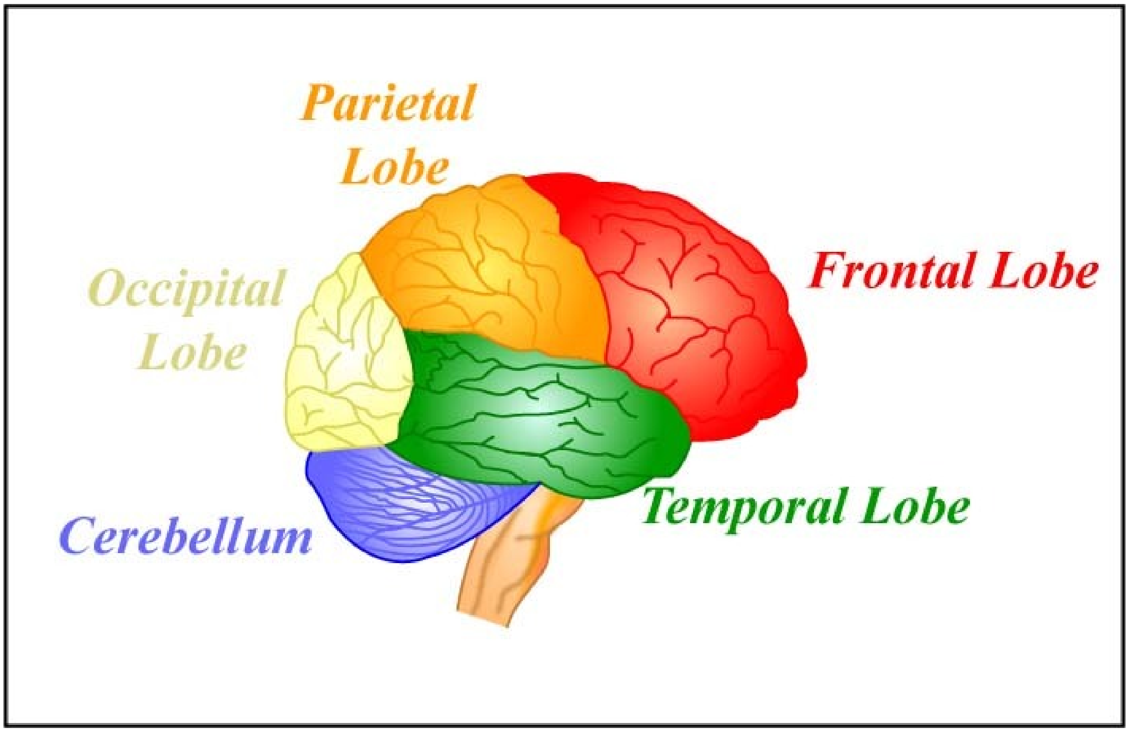
Figure 2. A sample of neuroanatomy nomenclature. The colored boxes indicate the different groupings of the seven structures printed in black, with the labels matching the color of the boxes. The hindbrain, midbrain, and forebrain nomenclature stems from the development of the vertebrate brain; these three areas differentiate early in embryonic development and later give rise to the structures listed in black. These three areas further subdivide into the telencephalon, diencephalon, mesencephalon, metencephalon, and myelencephalon at a later stage of development.

influence extends beyond that of movement and posture, given that it contains the greatest number of neurons of any structure in the brain. However, the exact role it plays in these higher functions is still a matter of further study.

Cerebral Hemispheres

The cerebral hemispheres are responsible for our cognitive abilities and conscious experience. They consist of the cerebral cortex and accompanying white matter ("cerebrum" in Latin) as well as the subcortical structures of the basal ganglia, amygdala, and hippocampal formation. The cerebral cortex is the largest and most visible part of the brain, retaining the Latin name (cerebrum) for "large brain" that Aristotle coined. It consists of two hemispheres (literally two half spheres) and gives the brain its characteristic gray and convoluted appearance; the folds and grooves of the cortex are called gyri and sulci (gyrus and sulcus if referring to just one), respectively.

The two cerebral hemispheres can be further subdivided into four lobes: the occipital, temporal, parietal, and frontal lobes. The occipital lobe is responsible for vision, as is much



The four lobes of the brain and the cerebellum. [Image: MIT OpenCourseWare, https://goo.gl/RwUEVt, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

of the temporal lobe. The temporal lobe is also involved in auditory processing, memory, and multisensory integration (e.g., the convergence of vision and audition). The parietal lobe houses the somatosensory (body sensations) cortex and structures involved in visual attention, as well as multisensory convergence zones. The frontal lobe houses the motor cortex and structures involved in motor planning, language, judgment, and decision-making. Not surprisingly then, the frontal lobe is proportionally larger in humans than in any other animal.

The subcortical structures are so named because they reside beneath the cortex. The basal ganglia are critical to voluntary movement and as such make contact with the cortex, the thalamus, and the brain stem. The amygdala and hippocampal formation are part of the limbic system, which also includes some cortical structures. The limbic system plays an important role in emotion and, in particular, in aversion and gratification.

A Brain Divided

The two cerebral hemispheres are connected by a dense bundle of white matter tracts called the corpus callosum. Some functions are replicated in the two hemispheres. For example, both hemispheres are responsible for sensory and motor function, although the sensory and motor cortices have a **contralateral** (or opposite-side) representation; that is, the left cerebral hemisphere is responsible for movements and sensations on the right side of the body and the right cerebral hemisphere is responsible for movements and sensations on the left side of the body. Other functions are **lateralized**; that is, they reside primarily in one hemisphere or the other. For example, for right-handed and the majority of left-handed individuals, the left hemisphere is most responsible for language.

There are some people whose two hemispheres are not connected, either because the corpus callosum was surgically severed (**callosotomy**) or due to a genetic abnormality. These **split-brain patients** have helped us understand the functioning of the two hemispheres. First, because of the contralateral representation of sensory information, if an object is placed in only the left or only the right **visual hemifield**, then only the right or left hemisphere, respectively, of the split-brain patient will see it. In essence, it is as though the person has two brains in his or her head, each seeing half the world. Interestingly, because language is very often localized in the left hemisphere, if we show the right hemisphere a picture and ask the patient what she saw, she will say she didn't see anything (because only the left hemisphere can speak and it didn't see anything). However, we know that the right hemisphere sees the picture because if the patient is asked to press a button whenever she sees the image, the left hand (which is controlled by the right hemisphere) will respond despite the left hemisphere's denial that anything was there. There are also some advantages to having disconnected hemispheres. Unlike those with a fully functional corpus callosum, a split-brain patient can simultaneously search for something in his right and left visual fields (Luck, Hillyard, Mangun, & Gazzaniga, 1989) and can do the equivalent of rubbing his stomach and patting his head at the same time (Franz, Eliason, Ivry, & Gazzaniga, 1996). In other words, they exhibit less competition between the hemispheres.

Gray Versus White Matter

The cerebral hemispheres contain both grey and white matter, so called because they appear grayish and whitish in dissections or in an MRI (magnetic resonance imaging; see, "Studying the Human Brain"). The **gray matter** is composed of the neuronal cell bodies (see module, "Neurons"). The cell bodies (or soma) contain the genes of the cell and are responsible for metabolism (keeping the cell alive) and synthesizing proteins. In this way, the cell body is the workhorse of the cell. The **white matter** is composed of the axons of the neurons, and, in particular, axons that are covered with a sheath of **myelin** (fatty support cells that are whitish in color). Axons conduct the electrical signals from the cell and are, therefore, critical to cell communication. People use the expression "use your gray matter" when they want a person to think harder. The "gray matter" in this expression is probably a reference to the cerebral hemispheres more generally; the gray cortical sheet (the convoluted surface of the cortex) being the most visible. However, both the gray matter and white matter are critical to proper functioning of the mind. Losses of either result in deficits in language, memory, reasoning, and other mental functions. See Figure 3 for MRI slices showing both the inner white matter that connects the cell bodies in the gray cortical sheet.

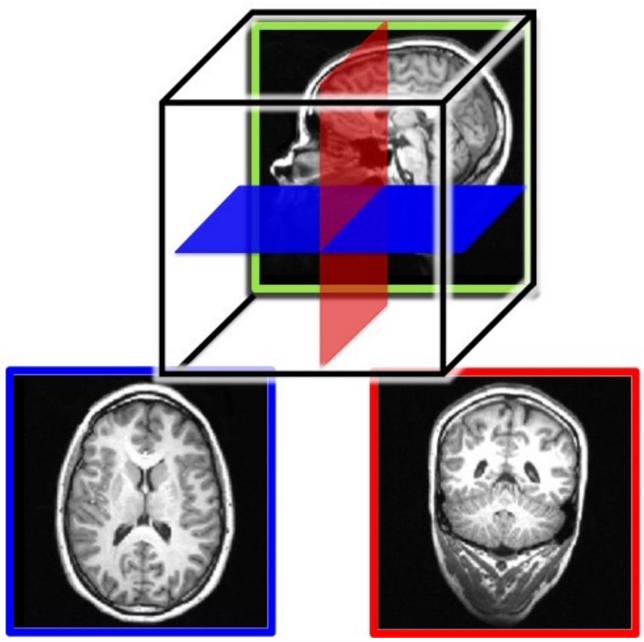


Figure 3. MRI slices of the human brain. Both the outer gray matter and inner white matter are visible in each image. The brain is a three-dimensional (3-D) structure, but an image is two-dimensional (2-D). Here, we show example slices of the three possible 2-D cuts through the brain: a saggital slice (top image), a horizontal slice (bottom left), which is also know as a transverse or axial slice, and a coronal slice (bottom right). The bottom two images are color coded to match the illustration of the relative orientations of the three slices in the top image.

Studying the Human Brain

How do we know what the brain does? We have gathered knowledge about the functions of the brain from many different methods. Each method is useful for answering distinct types of questions, but the strongest evidence for a specific role or function of a particular brain area is converging evidence; that is, similar findings reported from multiple studies using different methods.

One of the first organized attempts to study the functions of the brain was phrenology, a popular field of study in the first half of the 19th century. Phrenologists assumed that various features of the brain, such as its uneven surface, are reflected on the skull; therefore, they attempted to correlate bumps and indentations of the skull with specific functions of the brain. For example, they would claim that a very artistic person has ridges on the head that vary in size and location from those of someone who is very good at spatial reasoning. Although the assumption that the skull reflects the underlying brain structure has been proven wrong, phrenology nonetheless significantly impacted current-day neuroscience and its thinking about the functions of the brain. That is, different parts of the brain are devoted to very specific functions that can be identified through scientific inquiry.

Neuroanatomy

Dissection of the brain, in either animals or cadavers, has been a critical tool of neuroscientists since 340 BC when Aristotle first published his dissections. Since then this method has advanced considerably with the discovery of various staining techniques that can highlight particular cells. Because the brain can be sliced very thinly, examined under the microscope, and particular cells highlighted, this method is especially useful for studying specific groups of neurons or small brain structures; that is, it has a very high spatial resolution. Dissections allow scientists to study changes in the brain that occur due to various diseases or experiences (e.g., exposure to drugs or brain injuries).

Virtual dissection studies with living humans are also conducted. Here, the brain is imaged using computerized axial tomography (CAT) or MRI scanners; they reveal with very high precision the various structures in the brain and can help detect changes in gray or white matter. These changes in the brain can then be correlated with behavior, such as performance on memory tests, and, therefore, implicate specific brain areas in certain cognitive functions.

Changing the Brain

Some researchers induce lesions or ablate (i.e., remove) parts of the brain in animals. If the animal's behavior changes after the lesion, we can infer that the removed structure is important for that behavior. Lesions of human brains are studied in patient populations only; that is, patients who have lost a brain region due to a stroke or other injury, or who have had surgical removal of a structure to treat a particular disease (e.g., a callosotomy to control epilepsy, as in split-brain patients). From such case studies, we can infer brain function by measuring changes in the behavior of the patients before and after the lesion.

Because the brain works by generating electrical signals, it is also possible to change brain function with electrical stimulation. Transcranial magnetic stimulation (TMS) refers to a technique whereby a brief magnetic pulse is applied to the head that temporarily induces a weak electrical current in the brain. Although effects of TMS are sometimes referred to as temporary virtual lesions, it is more appropriate to describe the induced electricity as interference with neurons' normal communication with each other. TMS allows very precise study of when events in the brain happen so it has a good temporal resolution, but its application is limited only to the surface of the cortex and cannot extend to deep areas of the brain.

Transcranial direct current stimulation (tDCS) is similar to TMS except that it uses electrical current directly, rather than inducing it with magnetic pulses, by placing small electrodes on the skull. A brain area is stimulated by a low current (equivalent to an AA battery) for a more extended period of time than TMS. When used in combination with cognitive training, tDCS has been shown to improve performance of many cognitive functions such as mathematical ability, memory, attention, and coordination (e.g., Brasil-Neto, 2012; Feng, Bowden, & Kautz, 2013; Kuo & Nitsche, 2012).

Neuroimaging

Neuroimaging tools are used to study the brain in action; that is, when it is engaged in a specific task. **Positron emission tomography (PET)** records blood flow in the brain. The PET scanner detects the radioactive substance that is injected into the bloodstream of the participant just before or while he or she is performing some task (e.g., adding numbers). Because active neuron populations require metabolites, more blood and hence more radioactive substance flows into those regions. PET scanners detect the injected radioactive substance in specific brain regions, allowing researchers to infer that those areas were active during the task. **Functional magnetic resonance imaging (fMRI)** also relies on blood flow in the brain. This method, however, measures the changes in oxygen levels in the blood and does not require any substance to be injected into the participant. Both of these tools have good spatial resolution (although not as precise as dissection studies), but because it takes at least several seconds for the blood to arrive to the active areas of the brain, PET and fMRI have poor temporal resolution; that is, they do not tell us very precisely when the activity

A researcher looking at the areas of activation in the brain of a study participant who had an fMRI scan – areas of brain activation are determined by the amount of blood flow to a certain area – the more blood flow, the higher the activation of that area of the brain. [Image: National Institute of Mental Health, CC0 Public Domain, https://goo.gl/m25gce] occurred.

Electroencephalography (EEG), on the other hand, measures the electrical activity of the brain, and therefore, it has a much greater temporal resolution (millisecond precision rather than seconds) than PET or fMRI. Like tDCS, electrodes are placed on the participant's head when he or she is performing a task. In this case, however, many more electrodes are used, and they measure rather than produce activity. Because the electrical activity picked up at any particular electrode can be coming from anywhere in the brain, EEG has poor spatial resolution; that is, we have only a rough idea of which part of the brain generates the measured activity.

Diffuse optical imaging (DOI) can give researchers the best of both worlds: high spatial and temporal resolution, depending on how it is used. Here, one shines infrared light into the brain, and measures the light that comes back out. DOI relies on the fact that the properties of the light change when it passes through oxygenated blood, or when it encounters active neurons. Researchers can then infer from the properties of the collected light what regions in the brain were engaged by the task. When DOI is set up to detect changes in blood oxygen levels, the temporal resolution is low and comparable to PET or fMRI. However, when DOI is set up to directly detect active neurons, it has both high spatial and temporal resolution.

Because the spatial and temporal resolution of each tool varies, strongest evidence for what role a certain brain area serves comes from converging evidence. For example, we are more likely to believe that the hippocampal formation is involved in memory if multiple studies using a variety of tasks and different neuroimaging tools provide evidence for this hypothesis. The brain is a complex system, and only advances in brain research will show whether the brain can ever really understand itself.

Outside Resources

Video: Brain Bank at Harvard (National Geographic video)
http://video.nationalgeographic.com/video/science/health-human-body-sci/human-body/brain-bank-sci/

Video: Videos of a split-brain patient
http://youtu.be/ZMLzP1VCANo

Web: Atlas of the Human Brain: interactive demos and brain sections
http://www.thehumanbrain.info/

Web: Harvard University Human Brain Atlas: normal and diseased brain scans
http://www.med.harvard.edu/aanlib/home.html

Discussion Questions

1. In what ways does the segmentation of the brain into the brain stem, cerebellum, and cerebral hemispheres provide a natural division?
2. How has the study of split-brain patients been informative?
3. What is behind the expression "use your gray matter," and why is it not entirely accurate?
4. Why is converging evidence the best kind of evidence in the study of brain function?
5. If you were interested in whether a particular brain area was involved in a specific behavior, what neuroscience methods could you use?
6. If you were interested in the precise time in which a particular brain process occurred, which neuroscience methods could you use?

Vocabulary

Ablation
Surgical removal of brain tissue.

Axial plane
See “horizontal plane.”

Basal ganglia
Subcortical structures of the cerebral hemispheres involved in voluntary movement.

Brain stem
The “trunk” of the brain comprised of the medulla, pons, midbrain, and diencephalon.

Callosotomy
Surgical procedure in which the corpus callosum is severed (used to control severe epilepsy).

Case study
A thorough study of a patient (or a few patients) with naturally occurring lesions.

Cerebellum
The distinctive structure at the back of the brain, Latin for “small brain.”

Cerebral cortex
The outermost gray matter of the cerebrum; the distinctive convoluti characteristic of the mammalian brain.

Cerebral hemispheres
The cerebral cortex, underlying white matter, and subcortical structures.

Cerebrum
Usually refers to the cerebral cortex and associated white matter, but in some texts includes the subcortical structures.

Contralateral
Literally “opposite side”; used to refer to the fact that the two hemispheres of the brain process sensory information and motor commands for the opposite side of the body (e.g., the left hemisphere controls the right side of the body).

Converging evidence
Similar findings reported from multiple studies using different methods.

Coronal plane
A slice that runs from head to foot; brain slices in this plane are similar to slices of a loaf of bread, with the eyes being the front of the loaf.

Diffuse optical imaging (DOI)
A neuroimaging technique that infers brain activity by measuring changes in light as it is passed through the skull and surface of the brain.

Electroencephalography (EEG)
A neuroimaging technique that measures electrical brain activity via multiple electrodes on the scalp.

Frontal lobe
The front most (anterior) part of the cerebrum; anterior to the central sulcus and responsible for motor output and planning, language, judgment, and decision-making.

Functional magnetic resonance imaging (fMRI)
Functional magnetic resonance imaging (fMRI): A neuroimaging technique that infers brain activity by measuring changes in oxygen levels in the blood.

Gray matter
The outer grayish regions of the brain comprised of the neurons' cell bodies.

Gyri
(plural) Folds between sulci in the cortex.

Gyrus
A fold between sulci in the cortex.

Horizontal plane
A slice that runs horizontally through a standing person (i.e., parallel to the floor); slices of brain in this plane divide the top and bottom parts of the brain; this plane is similar to slicing a hamburger bun.

Lateralized
To the side; used to refer to the fact that specific functions may reside primarily in one

hemisphere or the other (e.g., for the majority individuals, the left hemisphere is most responsible for language).

Lesion
A region in the brain that suffered damage through injury, disease, or medical intervention.

Limbic system
Includes the subcortical structures of the amygdala and hippocampal formation as well as some cortical structures; responsible for aversion and gratification.

Metabolite
A substance necessary for a living organism to maintain life.

Motor cortex
Region of the frontal lobe responsible for voluntary movement; the motor cortex has a contralateral representation of the human body.

Myelin
Fatty tissue, produced by glial cells (see module, "Neurons") that insulates the axons of the neurons; myelin is necessary for normal conduction of electrical impulses among neurons.

Nomenclature
Naming conventions.

Occipital lobe
The back most (posterior) part of the cerebrum; involved in vision.

Parietal lobe
The part of the cerebrum between the frontal and occipital lobes; involved in bodily sensations, visual attention, and integrating the senses.

Phrenology
A now-discredited field of brain study, popular in the first half of the 19th century that correlated bumps and indentations of the skull with specific functions of the brain.

Positron emission tomography (PET)
A neuroimaging technique that measures brain activity by detecting the presence of a radioactive substance in the brain that is initially injected into the bloodstream and then pulled in by active brain tissue.

Sagittal plane
A slice that runs vertically from front to back; slices of brain in this plane divide the left and right side of the brain; this plane is similar to slicing a baked potato lengthwise.

Somatosensory (body sensations) cortex
The region of the parietal lobe responsible for bodily sensations; the somatosensory cortex has a contralateral representation of the human body.

Spatial resolution
A term that refers to how small the elements of an image are; high spatial resolution means the device or technique can resolve very small elements; in neuroscience it describes how small of a structure in the brain can be imaged.

Split-brain patient
A patient who has had most or all of his or her corpus callosum severed.

Subcortical
Structures that lie beneath the cerebral cortex, but above the brain stem.

Sulci
(plural) Grooves separating folds of the cortex.

Sulcus
A groove separating folds of the cortex.

Temporal lobe
The part of the cerebrum in front of (anterior to) the occipital lobe and below the lateral fissure; involved in vision, auditory processing, memory, and integrating vision and audition.

Temporal resolution
A term that refers to how small a unit of time can be measured; high temporal resolution means capable of resolving very small units of time; in neuroscience it describes how precisely in time a process can be measured in the brain.

Transcranial direct current stimulation (tDCS)
A neuroscience technique that passes mild electrical current directly through a brain area by placing small electrodes on the skull.

Transcranial magnetic stimulation (TMS)

A neuroscience technique whereby a brief magnetic pulse is applied to the head that temporarily induces a weak electrical current that interferes with ongoing activity.

Transverse plane
See “horizontal plane.”

Visual hemifield
The half of visual space (what we see) on one side of fixation (where we are looking); the left hemisphere is responsible for the right visual hemifield, and the right hemisphere is responsible for the left visual hemifield.

White matter
The inner whitish regions of the cerebrum comprised of the myelinated axons of neurons in the cerebral cortex.

References

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, 49, 1154–1165.

Brasil-Neto, J. P. (2012). Learning, memory, and transcranial direct current stimulation. *Frontiers in Psychiatry*, 3(80). doi: 10.3389/fpsyt.2012.00080.

Feng, W. W., Bowden, M. G., & Kautz, S. (2013). Review of transcranial direct current stimulation in poststroke recovery. *Topics in Stroke Rehabilitation*, 20, 68–77.

Franz, E. A., Eliassen, J. C., Ivry, R. B., & Gazzaniga, M. S. (1996). Dissociation of spatial and temporal coupling in the bimanual movements of callosotomy patients. *Psychological Science*, 7, 306–310.

Kandal, E. R., Schwartz, J. H., & Jessell, T. M. (Eds.) (2000). *Principles of neural science* (Vol. 4). New York, NY: McGraw-Hill.

Kuo, M. F., & Nitsche, M. A. (2012). Effects of transcranial electrical stimulation on cognition. *Clinical EEG and Neuroscience*, 43, 192–199.

Luck, S. J., Hillyard, S. A., Mangun, G. R., & Gazzaniga, M. S. (1989). Independent hemispheric attentional systems mediate visual search in split-brain patients. *Nature*, 342, 543–545.

Swanson, L. (2000). What is the brain? *Trends in Neurosciences*, 23, 519–527.

6
The Nervous System

Aneeq Ahmad

The mammalian nervous system is a complex biological organ, which enables many animals including humans to function in a coordinated fashion. The original design of this system is preserved across many animals through evolution; thus, adaptive physiological and behavioral functions are similar across many animal species. Comparative study of physiological functioning in the nervous systems of different animals lend insights to their behavior and their mental processing and make it easier for us to understand the human brain and behavior. In addition, studying the development of the nervous system in a growing human provides a wealth of information about the change in its form and behaviors that result from this change. The nervous system is divided into central and peripheral nervous systems, and the two heavily interact with one another. The peripheral nervous system controls volitional (somatic nervous system) and nonvolitional (autonomic nervous system) behaviors using cranial and spinal nerves. The central nervous system is divided into forebrain, midbrain, and hindbrain, and each division performs a variety of tasks; for example, the cerebral cortex in the forebrain houses sensory, motor, and associative areas that gather sensory information, process information for perception and memory, and produce responses based on incoming and inherent information. To study the nervous system, a number of methods have evolved over time; these methods include examining brain lesions, microscopy, electrophysiology, electroencephalography, and many scanning technologies.

Learning Objectives

- Describe and understand the development of the nervous system.
- Learn and understand the two important parts of the nervous system.
- Explain the two systems in the peripheral nervous system and what you know about the different regions and areas of the central nervous system.

- Learn and describe different techniques of studying the nervous system. Understand which of these techniques are important for cognitive neuroscientists.
- Describe the reasons for studying different nervous systems in animals other than human beings. Explain what lessons we learn from the evolutionary history of this organ.

Evolution of the Nervous System

Many scientists and thinkers (Cajal, 1937; Crick & Koch, 1990; Edelman, 2004) believe that the human nervous system is the most complex machine known to man. Its complexity points to one undeniable fact—that it has evolved slowly over time from simpler forms. Evolution of the nervous system is intriguing not because we can marvel at this complicated biological structure, but it is fascinating because it inherits a lineage of a long history of many less complex nervous systems (Figure 1), and it documents a record of adaptive behaviors observed in life forms other than humans. Thus, evolutionary study of the nervous system is important, and it is the first step in understanding its design, its workings, and its functional interface with the environment.

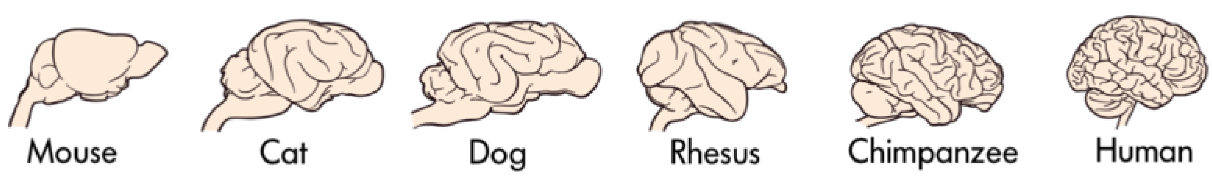


Figure 1 The brains of various animals

The brains of some animals, like apes, monkeys, and rodents, are structurally similar to humans (Figure 1), while others are not (e.g., invertebrates, single-celled organisms). Does anatomical similarity of these brains suggest that behaviors that emerge in these species are also similar? Indeed, many animals display behaviors that are similar to humans, e.g., apes use nonverbal communication signals with their hands and arms that resemble nonverbal forms of communication in humans (Gardner & Gardner, 1969; Goodall, 1986; Knapp & Hall, 2009). If we study very simple behaviors, like physiological responses made by individual neurons, then brain-based behaviors of invertebrates (Kandel & Schwartz, 1982) look very similar to humans, suggesting that from time immemorial such basic behaviors have been conserved in the brains of many simple animal forms and in fact are the foundation of more complex behaviors in animals that evolved later (Bullock, 1984).

Even at the micro-anatomical level, we note that individual neurons differ in complexity across animal species. Human neurons exhibit more intricate complexity than other animals; for example, neuronal processes (dendrites) in humans have many more branch points, branches, and spines.

Complexity in the structure of the nervous system, both at the macro- and micro-levels, give rise to complex behaviors. We can observe similar movements of the limbs, as in nonverbal communication, in apes and humans, but the variety and intricacy of nonverbal behaviors using hands in humans surpasses apes. Deaf individuals who use American Sign Language (ASL) express themselves in English nonverbally; they use this language with such fine gradation that many accents of ASL exist (Walker, 1987). Complexity of behavior with increasing complexity of the nervous system, especially the cerebral cortex, can be observed in the genus *Homo* (Figure 2). If we compare sophistication of material culture in Homo habilis (2 million years ago; brain volume ~650 cm3) and Homo sapiens (300,000 years to now; brain volume ~1400 cm3), the evidence shows that *Homo habilis* used crude stone tools compared with modern tools used by *Homo sapiens* to erect cities, develop written languages, embark on space travel, and study her own self. All of this is due to increasing complexity of the nervous system.

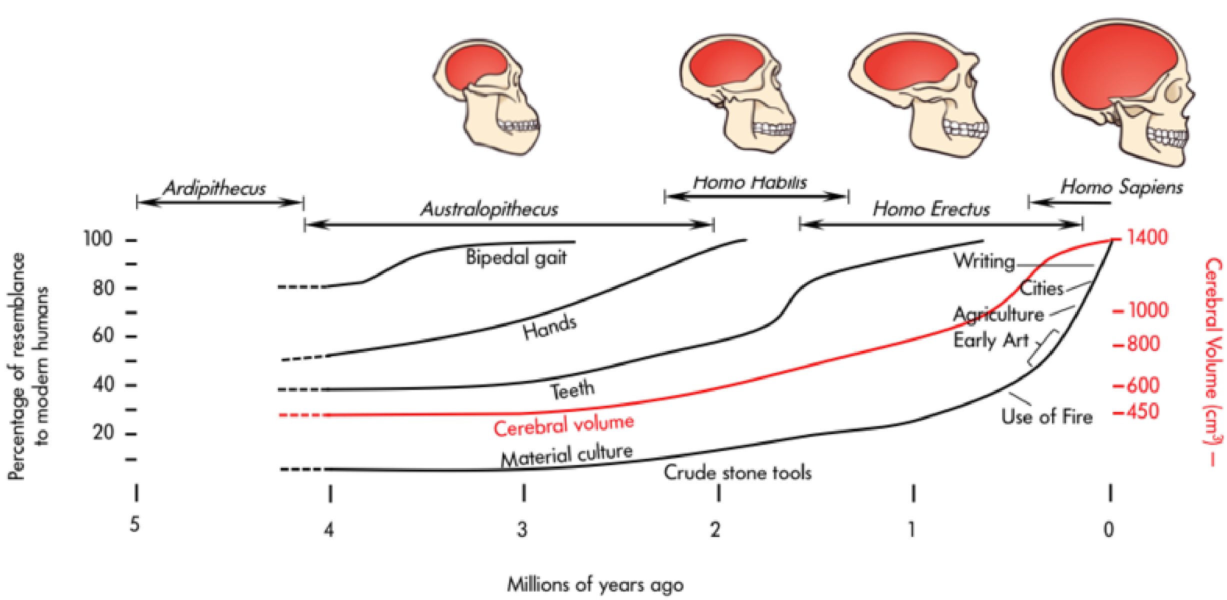


Figure 2 Changes in cerebral volume across evolution

What has led to the complexity of the brain and nervous system through evolution, to its behavioral and cognitive refinement? Darwin (1859, 1871) proposed two forces of natural and sexual selection as work engines behind this change. He prophesied, "psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation" that is, psychology will be based on evolution (Rosenzweig, Breedlove, & Leiman, 2002).

Development of the Nervous System

Where the study of change in the nervous system over eons is immensely captivating, studying the change in a single brain during individual development is no less engaging. In many ways the ontogeny (development) of the nervous system in an individual mimics the evolutionary advancement of this structure observed across many animal species. During development, the nervous tissue emerges from the **ectoderm** (one of the three layers of the mammalian embryo) through the process of **neural induction**. This process causes the formation of the neural tube, which extends in a **rostrocaudal** (head-to-tail) plane. The tube, which is hollow, seams itself in the rostrocaudal direction. In some disease conditions, the neural tube does not close caudally and results in an abnormality called **spina bifida**. In this pathological condition, the lumbar and sacral segments of the spinal cord are disrupted.

As gestation progresses, the neural tube balloons up (cephalization) at the rostral end, and **forebrain**, midbrain, hindbrain, and the spinal cord can be visually delineated (day 40). About 50 days into gestation, six cephalic areas can be anatomically discerned (also see below for a more detailed description of these areas).

The progenitor cells (**neuroblasts**) that form the lining (**neuroepithelium)** of the neural tube generate all the neurons and glial cells of the central nervous system. During early stages of this development, neuroblasts rapidly divide and specialize into many varieties of neurons and glial cells, but this proliferation of cells is not uniform along the neural tube—that is why we see the forebrain and hindbrain expand into larger cephalic tissues than the midbrain. The neuroepithelium also generates a group of specialized cells that migrate outside the neural tube to form the **neural crest**. This structure gives rise to sensory and autonomic neurons in the peripheral nervous system.

The Structure of the Nervous System

The mammalian nervous system is divided into central and peripheral nervous systems.

The Peripheral Nervous System

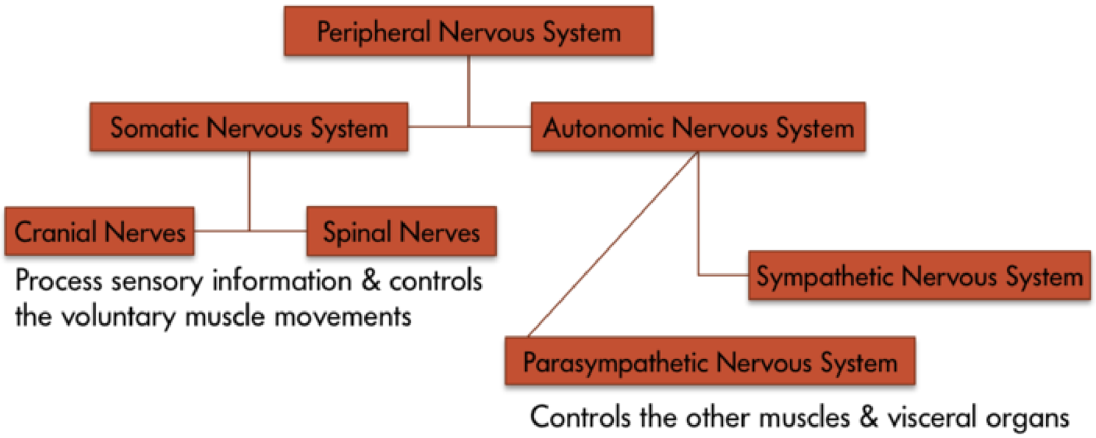


Figure 3 The various components of the peripheral nervous system

The peripheral nervous system is divided into somatic and autonomic nervous systems (Figure 3). Where the somatic nervous system consists of cranial nerves (12 pairs) and spinal nerves (31 pairs) and is under the volitional control of the individual in maneuvering bodily muscles, the autonomic nervous system also running through these nerves lets the individual have little control over muscles and glands. Main divisions of the autonomic nervous system that control visceral structures are the sympathetic and parasympathetic nervous systems.

At an appropriate cue (say a fear-inducing object like a snake), the sympathetic division generally energizes many muscles (e.g., heart) and glands (e.g., adrenals), causing activity and release of hormones that lead the individual to negotiate the fear-causing snake with fight-or-flight responses. Whether the individual decides to *fight* the snake or *run* away from it, either action requires energy; in short, the sympathetic nervous system says "go, go, go." The parasympathetic nervous system, on the other hand, curtails undue energy mobilization into muscles and glands and modulates the response by saying "stop, stop, stop." This push–

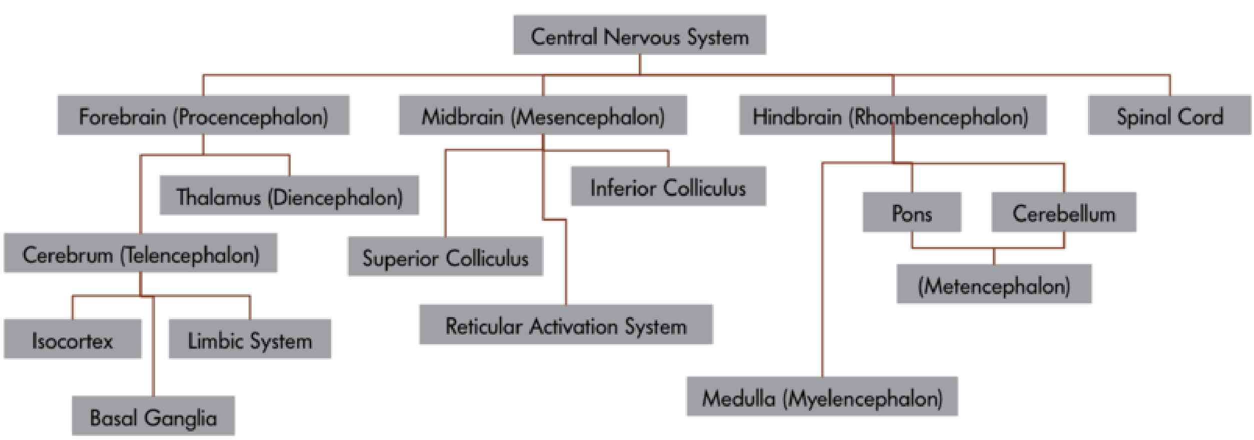


Figure 4 the central nervous system and its components pull tandem system regulates fight-or-flight responses in all of us.

The Central Nervous System

The central nervous system is divided into a number of important parts (see Figure 4), including the spinal cord, each specialized to perform a set of specific functions. Telencephalon or cerebrum is a *newer* development in the evolution of the mammalian nervous system. In humans, it is about the size of a large napkin and when crumpled into the skull, it forms furrows called sulci (singular form, sulcus). The bulges between sulci are called gyri (singular form, gyrus). The cortex is divided into two hemispheres, and each hemisphere is further divided into four lobes (Figure 5a), which have specific functions. The division of these lobes is based on two delineating sulci: the central sulcus divides the hemisphere into frontal and parietal-occipital lobes and the lateral sulcus marks the temporal lobe, which lies below.

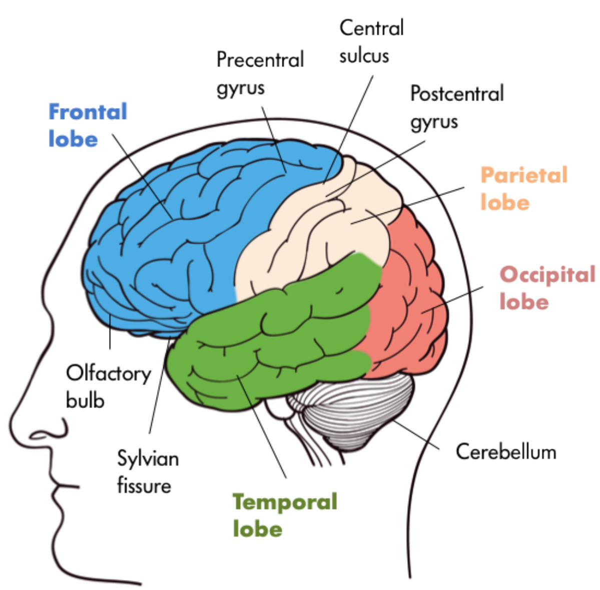


Figure 5a The lobes of the brain

Just in front of the central sulcus lies an area called the primary motor cortex (precentral gyrus), which connects to the muscles of the body, and on volitional command moves them. From mastication to movements in the genitalia, the body map is represented on this strip (Figure 5b).

Some body parts, like fingers, thumbs, and lips, occupy a greater representation on the strip than, say, the trunk. This disproportionate representation of the body on the primary motor cortex is called the magnification factor (Rolls & Cowey, 1970) and is seen in other motor and sensory areas. At the lower end of the central sulcus, close to the lateral sulcus, lies the Broca's area (Figure 6b) in the left frontal lobe, which is involved with language production. Damage to this part of the brain led Pierre Paul Broca, a French neuroscientist in 1861, to document many different forms of aphasias, in which his patients would lose the ability to speak or would retain partial speech impoverished in syntax and grammar (AAAS, 1880). It is no wonder that others have found subvocal rehearsal and central executive processes of working memory in this frontal lobe (Smith & Jonides, 1997, 1999).

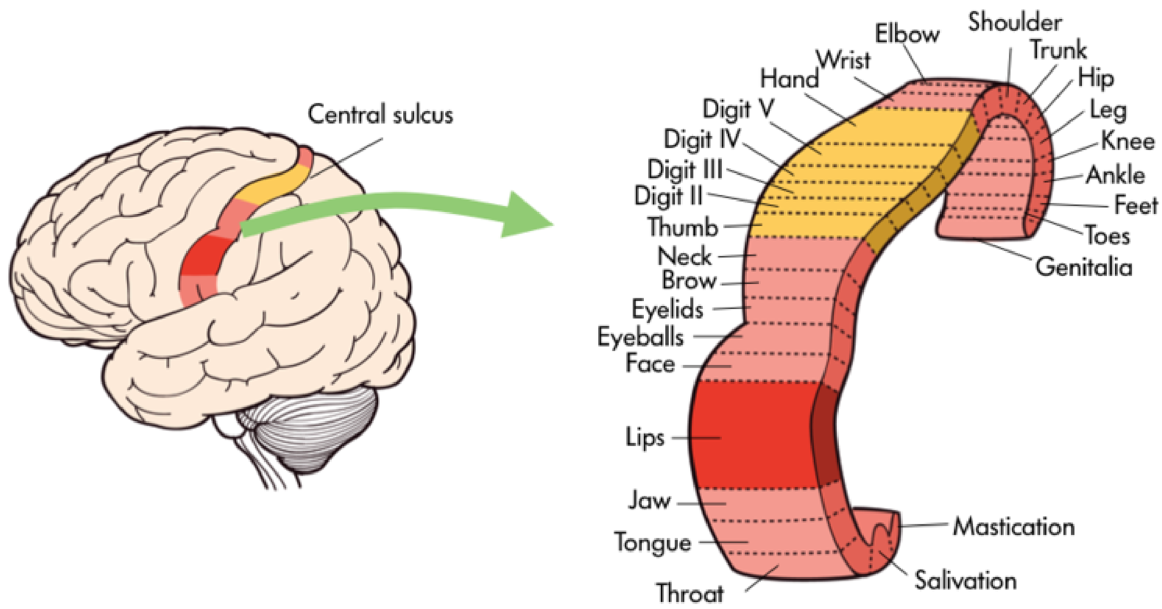


Figure 5b. Specific body parts like the tongue or fingers are mapped onto certain areas of the brain including the primary motor cortex.

Just behind the central gyrus, in the **parietal lobe**, lies the **primary somatosensory cortex** (Figure 6a) on the postcentral gyrus, which represents the whole body receiving inputs from the skin and muscles. The primary somatosensory cortex parallels, abuts, and connects heavily to the primary motor cortex and resembles it in terms of areas devoted to bodily representation. All spinal and some cranial nerves (e.g., the facial nerve) send sensory signals from skin (e.g., touch) and muscles to the primary somatosensory cortex. Close to the lower (ventral) end of this strip, curved inside the parietal lobe, is the taste area (secondary somatosensory cortex), which is involved with taste experiences that originate from the tongue, pharynx, epiglottis, and so forth.

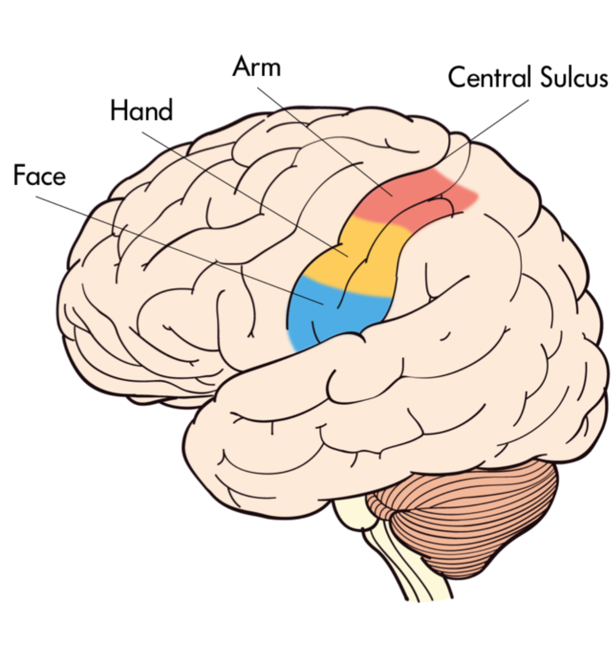


Figure 6a The Primary Somatosensory Cortex

Just below the parietal lobe, and under the caudal end of the lateral fissure, in the temporal lobe, lies the Wernicke's area (Demonet et al., 1992). This area is involved with language comprehension and is connected to the Broca's area through the arcuate fasciculus, nerve fibers that connect these two regions. Damage to the Wernicke's area (Figure 6b) results in many kinds of agnosias; agnosia is defined as an inability to know or understand language and speech-related behaviors. So an individual may show word deafness, which is an inability to recognize spoken language, or word blindness, which is an inability to recognize written or printed language. Close in proximity to the Wernicke's area is the primary auditory cortex, which is involved with audition, and finally the brain region devoted to smell (olfaction) is tucked away inside the primary olfactory cortex (prepyriform cortex).

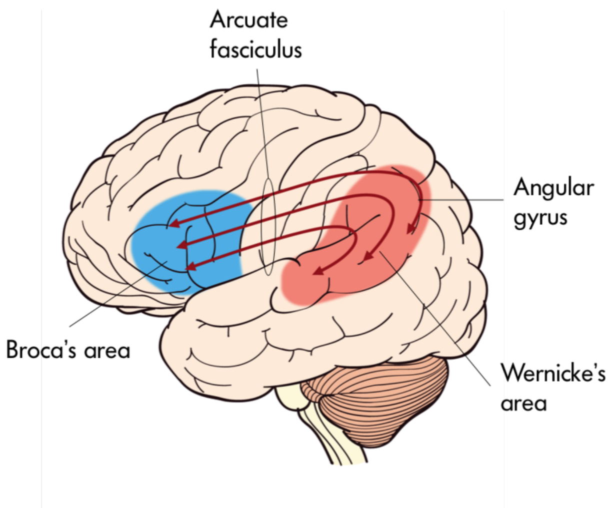


Figure 6b Wernicke's area

At the very back of the cerebral cortex lies the occipital lobe housing the primary visual cortex. Optic nerves travel all the way to the thalamus (lateral geniculate nucleus, LGN) and then to visual cortex, where images that are received on the retina are projected (Hubel, 1995).

In the past 50 to 60 years, visual sense and visual pathways have been studied extensively, and our understanding about them has increased manifold. We now understand that all objects that form images on the retina are transformed (transduction) in neural language handed down to the visual cortex for further processing. In the visual cortex, all attributes (features) of the image, such as the color, texture, and orientation, are decomposed and processed by different visual cortical modules (Van Essen, Anderson & Felleman, 1992) and then recombined to give rise to singular perception of the image in question.

If we cut the cerebral hemispheres in the middle, a new set of structures come into view. Many of these perform different functions vital to our being. For example, the limbic system contains a number of nuclei that process memory (hippocampus and fornix) and attention and emotions (cingulate gyrus); the globus pallidus is involved with motor movements and their coordination; the hypothalamus and thalamus are involved with drives, motivations, and trafficking of sensory and motor throughputs. The hypothalamus plays a key role in regulating endocrine hormones in conjunction with the pituitary gland that extends from the hypothalamus through a stalk (infundibulum).

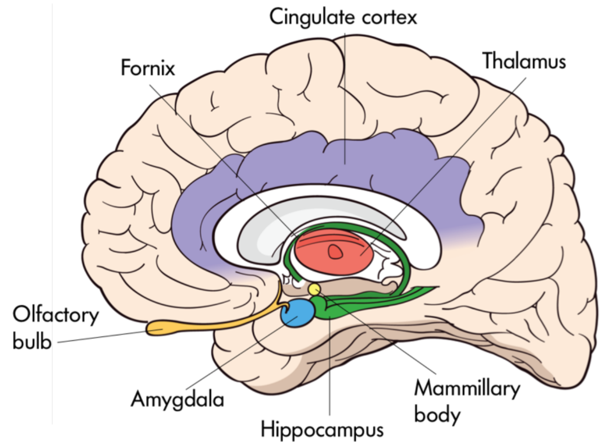


Figure 7 The interior of the brain

As we descend down the thalamus, the midbrain comes into view with superior and inferior colliculi, which process visual and auditory information, as does the substantia nigra, which is involved with notorious Parkinson's disease, and the reticular formation regulating arousal, sleep, and temperature. A little lower, the hindbrain with the **pons** processes sensory and motor information employing the cranial nerves, works as a bridge that connects the cerebral cortex with the medulla, and reciprocally transfers information back and forth between the brain and the spinal cord. The **medulla oblongata** processes breathing, digestion, heart and blood vessel function, swallowing, and sneezing. The **cerebellum** controls motor movement coordination, balance, equilibrium, and muscle tone.

The midbrain and the hindbrain, which make up the brain stem, culminate in the spinal cord. Whereas inside the cerebral cortex, the **gray matter** (neuronal cell bodies) lies outside and **white matter** (myelinated axons) inside; in the spinal cord this arrangement reverses, as the gray matter resides inside and the white matter outside. Paired nerves (ganglia) exit the spinal cord, some closer in direction towards the back (dorsal) and others towards the front (ventral). The dorsal nerves (afferent) receive sensory information from skin and muscles, and ventral nerves (efferent) send signals to muscles and organs to respond.

Studying the Nervous System

The study of the nervous system involves anatomical and physiological techniques that have improved over the years in efficiency and caliber. Clearly, gross morphology of the nervous system requires an eye-level view of the brain and the spinal cord. However, to resolve minute components, optical and electron microscopic techniques are needed.

Light microscopes and, later, electron microscopes have changed our understanding of the intricate connections that exist among nerve cells. For example, modern staining procedures (**immunocytochemistry**) make it possible to see selected neurons that are of one type or another or are affected by growth. With better resolution of the electron microscopes, fine structures like the synaptic cleft between the pre- and post-synaptic neurons can be studied in detail.

Along with the neuroanatomical techniques, a number of other methodologies aid neuroscientists in studying the function and physiology of the nervous system. Early on, **lesion studies** in animals (and study of neurological damage in humans) provided information about the function of the nervous system, by ablating (removing) parts of the nervous system or using neurotoxins to destroy them and documenting the effects on behavior or mental processes. Later, more sophisticated microelectrode techniques were introduced, which led to recording from single neurons in the animal brains and investigating their physiological functions. Such studies led to formulating theories about how sensory and motor information are processed in the brain. To study many neurons (millions of them at a time) electroencephalographic (EEG) techniques were introduced. These methods are used to study how large ensembles of neurons, representing different parts of the nervous system, with (**event-related potentials**) or without stimulation function together. In addition, many scanning techniques that visualize the brain in conjunction with methods mentioned above are used to understand the details of the structure and function of the brain. These include **computerized axial tomography** (CAT), which uses X-rays to capture many pictures of the brain and sandwiches them into 3-D models to study it. The resolution of this method is inferior to **magnetic resonance imaging** (MRI), which is yet another way to capture brain images using large magnets that bobble (precession) hydrogen nuclei in the brain. Although the resolution of MRI scans is much better than CAT scans, they do not provide any functional information about the brain. **Positron Emission Tomography** (PET) involves the acquisition of physiologic (functional) images of the brain based on the detection of positrons. Radio-labeled isotopes of certain chemicals, such as an analog of glucose (fluorodeoxyglucose), enters the active nerve cells and emits positrons, which are captured and mapped into scans. Such scans show how the brain and its many modules become active (or not) when energized with entering glucose analog. Disadvantages of PET scans include being invasive and rendering poor spatial resolution. The latter is why modern PET machines are coupled with CAT scanners to gain better resolution of the functioning brain. Finally, to avoid the invasiveness of PET, functional MRI (fMRI) techniques were developed. Brain images based on fMRI technique visualize brain function by changes in the flow of fluids (blood) in brain areas that occur over time. These scans provide a wealth of functional information about the brain as the individual may engage in a task, which is why the last two methods of brain scanning are very popular among cognitive neuroscientists.

Understanding the nervous system has been a long journey of inquiry, spanning several hundreds of years of meticulous studies carried out by some of the most creative and versatile investigators in the fields of philosophy, evolution, biology, physiology, anatomy, neurology, neuroscience, cognitive sciences, and psychology. Despite our profound understanding of this organ, its mysteries continue to surprise us, and its intricacies make us marvel at this complex structure unmatched in the universe.

Outside Resources

Video: Pt. 1 video on the anatomy of the nervous system
http://www.youtube.com/watch?v=D1zkVBHPh5c

Video: Pt. 2 video on the anatomy of the nervous system
http://www.youtube.com/watch?v=8hC6NGQReL4

Video: To look at functions of the brain and neurons, watch
http://www.youtube.com/watch?v=9UukcdU258A

Web: To look at different kinds of brains, visit
http://brainmuseum.org/

Discussion Questions

1. Why is it important to study the nervous system in an evolutionary context?
2. How can we compare changes in the nervous system made through evolution to changes made during development?
3. What are the similarities and differences between the somatic and autonomic nervous systems?
4. Describe functions of the midbrain and hindbrain.
5. Describe the anatomy and functions of the forebrain.
6. Compare and contrast electroencephalograms to electrophysiological techniques.
7. Which brain scan methodologies are important for cognitive scientists? Why?

Vocabulary

Afferent nerves
Nerves that carry messages to the brain or spinal cord.

Agnosias
Due to damage of Wernicke's area. An inability to recognize objects, words, or faces.

Aphasia
Due to damage of the Broca's area. An inability to produce or understand words.

Arcuate fasciculus
A fiber tract that connects Wernicke's and Broca's speech areas.

Autonomic nervous system
A part of the peripheral nervous system that connects to glands and smooth muscles. Consists of sympathetic and parasympathetic divisions.

Broca's area
An area in the frontal lobe of the left hemisphere. Implicated in language production.

Central sulcus
The major fissure that divides the frontal and the parietal lobes.

Cerebellum
A nervous system structure behind and below the cerebrum. Controls motor movement coordination, balance, equilibrium, and muscle tone.

Cerebrum
Consists of left and right hemispheres that sit at the top of the nervous system and engages in a variety of higher-order functions.

Cingulate gyrus
A medial cortical portion of the nervous tissue that is a part of the limbic system.

Computerized axial tomography
A noninvasive brain-scanning procedure that uses X-ray absorption around the head.

Ectoderm
The outermost layer of a developing fetus.

Efferent nerves
Nerves that carry messages from the brain to glands and organs in the periphery.

Electroencephalography
A technique that is used to measure gross electrical activity of the brain by placing electrodes on the scalp.

Event-related potentials
A physiological measure of large electrical change in the brain produced by sensory stimulation or motor responses.

Forebrain
A part of the nervous system that contains the cerebral hemispheres, thalamus, and hypothalamus.

Fornix
(plural form, fornices) A nerve fiber tract that connects the hippocampus to mammillary bodies.

Frontal lobe
The most forward region (close to forehead) of the cerebral hemispheres.

Functional magnetic resonance imaging
(or fMRI) A noninvasive brain-imaging technique that registers changes in blood flow in the brain during a given task (also see magnetic resonance imaging).

Globus pallidus
A nucleus of the basal ganglia.

Gray matter
Composes the bark or the cortex of the cerebrum and consists of the cell bodies of the neurons (see also white matter).

Gyrus
(plural form, gyri) A bulge that is raised between or among fissures of the convoluted brain.

Hippocampus
(plural form, hippocampi) A nucleus inside (medial) the temporal lobe implicated in learning and memory.

Homo habilis
A human ancestor, handy man, that lived two million years ago.

Homo sapiens
Modern man, the only surviving form of the genus Homo.

Hypothalamus
Part of the diencephalon. Regulates biological drives with pituitary gland.

Immunocytochemistry
A method of staining tissue including the brain, using antibodies.

Lateral geniculate nucleus
(or LGN) A nucleus in the thalamus that is innervated by the optic nerves and sends signals to the visual cortex in the occipital lobe.

Lateral sulcus
The major fissure that delineates the temporal lobe below the frontal and the parietal lobes.

Lesion studies
A surgical method in which a part of the animal brain is removed to study its effects on behavior or function.

Limbic system
A loosely defined network of nuclei in the brain involved with learning and emotion.

Magnetic resonance imaging
Or MRI is a brain imaging noninvasive technique that uses magnetic energy to generate brain images (also see fMRI).

Magnification factor
Cortical space projected by an area of sensory input (e.g., mm of cortex per degree of visual field).

Medulla oblongata

An area just above the spinal cord that processes breathing, digestion, heart and blood vessel function, swallowing, and sneezing.

Neural crest
A set of primordial neurons that migrate outside the neural tube and give rise to sensory and autonomic neurons in the peripheral nervous system.

Neural induction
A process that causes the formation of the neural tube.

Neuroblasts
Brain progenitor cells that asymmetrically divide into other neuroblasts or nerve cells.

Neuroepithelium
The lining of the neural tube.

Occipital lobe
The back part of the cerebrum, which houses the visual areas.

Parasympathetic nervous system
A division of the autonomic nervous system that is slower than its counterpart—that is, the sympathetic nervous system—and works in opposition to it. Generally engaged in "rest and digest" functions.

Parietal lobe
An area of the cerebrum just behind the central sulcus that is engaged with somatosensory and gustatory sensation.

Pons
A bridge that connects the cerebral cortex with the medulla, and reciprocally transfers information back and forth between the brain and the spinal cord.

Positron Emission Tomography
(or PET) An invasive procedure that captures brain images with positron emissions from the brain after the individual has been injected with radio-labeled isotopes.

Primary Motor Cortex
A strip of cortex just in front of the central sulcus that is involved with motor control.

Primary Somatosensory Cortex
A strip of cerebral tissue just behind the central sulcus engaged in sensory reception of bodily sensations.

Rostrocaudal
A front-back plane used to identify anatomical structures in the body and the brain.

Somatic nervous system
A part of the peripheral nervous system that uses cranial and spinal nerves in volitional actions.

Spina bifida
A developmental disease of the spinal cord, where the neural tube does not close caudally.

Sulcus
(plural form, sulci) The crevices or fissures formed by convolutions in the brain.

Sympathetic nervous system
A division of the autonomic nervous system, that is faster than its counterpart that is the parasympathetic nervous system and works in opposition to it. Generally engaged in “fight or flight” functions.

Temporal lobe
An area of the cerebrum that lies below the lateral sulcus; it contains auditory and olfactory (smell) projection regions.

Thalamus
A part of the diencephalon that works as a gateway for incoming and outgoing information.

Transduction
A process in which physical energy converts into neural energy.

Wernicke’s area
A language area in the temporal lobe where linguistic information is comprehended (Also see Broca’s area).

White matter
Regions of the nervous system that represent the axons of the nerve cells; whitish in color because of myelination of the nerve cells.

Working memory

Short transitory memory processed in the hippocampus.

References

American Association for the Advancement of Science (AAAS). (1880). Dr. Paul Broca. *Science*, 1(8), 93. http://www.jstor.org/stable/2900242

Bullock, T. H. (1984). Comparative neuroscience holds promise for quiet revolutions. *Science*, 225(4661), 473–478.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2, 263–275.

Darwin, C. (1871). *The descent of man, and the selection in relation to sex*. London: J. Murray.

Darwin, C. (1859). On the origins of species by means of natural selection, or, *The preservation of favoured races in the struggle for life*. London, UK: J. Murray.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., . . . Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115(6), 1753–1768.

Edelman, G. (2004). *Wider than the sky: The phenomenal gift of consciousness*. New Haven, CT: Yale University Press.

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165 (3894), 664–672.

Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Harvard University Press.

Hubel, D. H. (1995). *Eye, brain, and vision*. Freeman & Co., NY: Scientific American Library/ Scientific American Books.

Kandel, E. R., & Schwartz, J. H. (1982). Molecular biology of learning: Modulation of transmitter release. *Science*, 218(4571), 433–443.

Knapp, M. L., & Hall, J. A. (2009). *Nonverbal communication in human interaction*. Boston, MA: Wadsworth Cengage Learning.

Ramón y Cahal, S. (1937). Recollections of my life. *Memoirs of the American Philosophical Society*, 8, 1901–1917.

Rolls, E. T., & Cowey, A. (1970). Topography of the retina and striate cortex and its relationship to visual acuity in rhesus monkeys and squirrel monkeys. *Experimental Brain Research*, 10 (3), 298–310.

Rosenzweig, M. R., Breedlove, S. M., & Leiman, A. L. (2002). *Biological psychology* (3rd ed.). Sunderland, MA: Sinauer Associates.

Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, 283(5408), 1657–1661.

Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive psychology*, 33, 5–42.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science, 255*(5043), 419–423.

Walker, L. A. (1987). *A loss for words: The story of deafness in a family*. New York, NY: Harper Perennial.

7
Evolutionary Theories in Psychology

David M. Buss

Evolution or change over time occurs through the processes of natural and sexual selection. In response to problems in our environment, we adapt both physically and psychologically to ensure our survival and reproduction. Sexual selection theory describes how evolution has shaped us to provide a mating advantage rather than just a survival advantage and occurs through two distinct pathways: intrasexual competition and intersexual selection. Gene selection theory, the modern explanation behind evolutionary biology, occurs through the desire for gene replication. Evolutionary psychology connects evolutionary principles with modern psychology and focuses primarily on psychological adaptations: changes in the way we think in order to improve our survival. Two major evolutionary psychological theories are described: Sexual strategies theory describes the psychology of human mating strategies and the ways in which women and men differ in those strategies. Error management theory describes the evolution of biases in the way we think about everything.

Learning Objectives

- Learn what "evolution" means.
- Define the primary mechanisms by which evolution takes place.
- Identify the two major classes of adaptations.
- Define sexual selection and its two primary processes.
- Define gene selection theory.
- Understand psychological adaptations.
- Identify the core premises of sexual strategies theory.
- Identify the core premises of error management theory, and provide two empirical examples of adaptive cognitive biases.

Introduction

If you have ever been on a first date, you're probably familiar with the anxiety of trying to figure out what clothes to wear or what perfume or cologne to put on. In fact, you may even consider flossing your teeth for the first time all year. When considering why you put in all this work, you probably recognize that you're doing it to impress the other person. But how did you learn these particular behaviors? Where did you get the idea that a first date should be at a nice restaurant or someplace unique? It is possible that we have been taught these behaviors by observing others. It is also possible, however, that these behaviors—the fancy clothes, the expensive restaurant—are biologically programmed into us. That is, just as peacocks display their feathers to show how attractive they are, or some lizards do push-ups to show how strong they are, when we style our hair or bring a gift to a date, we're trying to communicate to the other person: "Hey, I'm a good mate! Choose me! Choose me!"

It may seem like just a casual date, but don't doubt that the forces of evolution are hard at work below the surface. [Image: Best Couples, https://goo.gl/aBMY6W, CC BY-SA 2.0, https://goo.gl/jSSrcO]

However, we all know that our ancestors hundreds of thousands of years ago weren't driving sports cars or wearing designer clothes to attract mates. So how could someone ever say that such behaviors are "biologically programmed" into us? Well, even though our ancestors might not have been doing these specific actions, these behaviors are the result of the same driving force: the powerful influence of **evolution**. Yes, evolution—certain traits and behaviors developing over time because they are advantageous to our survival. In the case of dating, doing something like offering a gift might represent more than a nice gesture. Just as chimpanzees will give food to mates to show they can provide for them, when you offer gifts to your dates, you are communicating that you have the money or "resources" to help take care of them. And even though the person receiving the gift may not realize it, the same evolutionary forces are influencing his or her behavior as well. The receiver of the gift evaluates not only the gift but also the gift-giver's clothes, physical appearance, and many other qualities, to determine whether the individual is a suitable mate. But because these evolutionary processes are hardwired into us, it is easy to overlook their influence.

To broaden your understanding of evolutionary processes, this module will present some of the most important elements of evolution as they impact psychology. Evolutionary theory helps us piece together the story of how we humans have prospered. It also helps to explain why we behave as we do on a daily basis in our modern world: why we bring gifts on dates, why we get jealous, why we crave our favorite foods, why we protect our children, and so on. Evolution may seem like a historical concept that applies only to our ancient ancestors but, in truth, it is still very much a part of our modern daily lives.

Basics of Evolutionary Theory

Evolution simply means change over time. Many think of evolution as the development of traits and behaviors that allow us to survive this "dog-eat-dog" world, like strong leg muscles to run fast, or fists to punch and defend ourselves. However, physical survival is only important if it eventually contributes to successful reproduction. That is, even if you live to be a 100-year-old, if you fail to mate and produce children, your genes will die with your body. Thus, *reproductive* success, not *survival* success, is the engine of evolution by **natural selection**. Every mating success by one person means the loss of a mating opportunity for another. Yet every living human being is an evolutionary success story. Each of us is descended from a long and unbroken line of ancestors who triumphed over others in the struggle to survive (at least long enough to mate) and reproduce. However, in order for our genes to endure over time—to survive harsh climates, to defeat predators—we have inherited adaptive, psychological processes designed to ensure success.

At the broadest level, we can think of organisms, including humans, as having two large classes of **adaptations**—or traits and behaviors that evolved over time to increase our reproductive success. The first class of adaptations are called survival adaptations: mechanisms that helped our ancestors handle the "hostile forces of nature." For example, in order to survive very hot temperatures, we developed sweat glands to cool ourselves. In order to survive very cold temperatures, we developed shivering mechanisms (the speedy contraction and expansion of muscles to produce warmth). Other examples of survival adaptations include developing a craving for fats and sugars, encouraging us to seek out particular foods rich in fats and sugars that keep us going longer during food shortages. Some threats, such as snakes, spiders, darkness, heights, and strangers, often produce fear in us, which encourages us to avoid them and thereby stay safe. These are also examples of survival adaptations. However, all of these adaptations are for physical *survival,* whereas the second class of adaptations are for *reproduction*, and help us compete for mates. These adaptations are described in an evolutionary theory proposed by Charles Darwin, called sexual selection theory.

Sexual Selection Theory

Darwin noticed that there were many traits and behaviors of organisms that could not be explained by “survival selection.” For example, the brilliant plumage of peacocks should actually lower their rates of survival. That is, the peacocks’ feathers act like a neon sign to predators, advertising “Easy, delicious dinner here!” But if these bright feathers only lower peacocks’ chances at survival, why do they have them? The same can be asked of similar characteristics of other animals, such as the large antlers of male stags or the wattles of roosters, which also seem to be unfavorable to survival. Again, if these traits only make the animals less likely to survive, why did they develop in the first place? And how have these animals continued to survive with these traits over thousands and thousands of years? Darwin’s answer to this conundrum was the theory of sexual selection: the evolution of characteristics, not because of survival advantage, but because of *mating* advantage.

Modern sports like boxing can be seen as modified/stylized versions of the evolutionary behavior of intrasexual competition. [Image: Dave Hogg, https://goo.gl/fL5U2Z, CC BY 2.0, https://goo.gl/9uSnqN]

Sexual selection occurs through two processes. The first, intrasexual competition, occurs when members of one sex compete against each other, and the winner gets to mate with a member of the opposite sex. Male stags, for example, battle with their antlers, and the winner (often the stronger one with larger antlers) gains mating access to the female. That is, even though large antlers make it harder for the stags to run through the forest and evade predators (which lowers their survival success), they provide the stags with a better chance of attracting a mate (which increases their reproductive success). Similarly, human males sometimes also compete against each other in physical contests: boxing, wrestling, karate, or group-on-group sports, such as football. Even though engaging in these activities poses a "threat" to their survival success, as with the stag, the victors are often more attractive to potential mates, increasing their reproductive success. Thus, whatever qualities lead to success in intrasexual competition are then passed on with greater frequency due to their association with greater mating success.

The second process of sexual selection is preferential mate choice, also called **intersexual selection**. In this process, if members of one sex are attracted to certain qualities in mates—such as brilliant plumage, signs of good health, or even intelligence—those desired qualities get passed on in greater numbers, simply because their possessors mate more often. For example, the colorful plumage of peacocks exists due to a long evolutionary history of peahens' (the term for female peacocks) attraction to males with brilliantly colored feathers.

In all sexually-reproducing species, adaptations in both sexes (males and females) exist due to survival selection and sexual selection. However, unlike other animals where one sex has dominant control over mate choice, humans have "mutual mate choice." That is, both women and men typically have a say in choosing their mates. And both mates value qualities such as kindness, intelligence, and dependability that are beneficial to long-term relationships—qualities that make good partners and good parents.

Gene Selection Theory

In modern evolutionary theory, all evolutionary processes boil down to an organism's genes. Genes are the basic "units of heredity," or the information that is passed along in DNA that tells the cells and molecules how to "build" the organism and how that organism should behave. Genes that are better able to encourage the organism to reproduce, and thus replicate themselves in the organism's offspring, have an advantage over competing genes that are less able. For example, take female sloths: In order to attract a mate, they will scream as loudly as they can, to let potential mates know where they are in the thick jungle. Now, consider two types of genes in female sloths: one gene that allows them to scream extremely loudly, and another that only allows them to scream moderately loudly. In this case, the sloth with the gene that allows her to shout louder will attract more mates—increasing reproductive success—which ensures that her genes are more readily passed on than those of the quieter sloth.

Essentially, genes can boost their own replicative success in two basic ways. First, they can influence the odds for survival and reproduction of the organism they are in (individual reproductive success or fitness—as in the example with the sloths). Second, genes can also influence the organism to help other organisms who also likely contain those genes—known as "genetic relatives"—to survive and reproduce (which is called inclusive fitness). For example, why do human parents tend to help their own kids with the financial burdens of a college education and not the kids next door? Well, having a college education increases one's attractiveness to other mates, which increases one's likelihood for reproducing and passing on genes. And because parents' genes are in their own children (and not the neighborhood children), funding their children's educations increases the likelihood that the parents' genes will be passed on.

Understanding gene replication is the key to understanding modern evolutionary theory. It also fits well with many evolutionary psychological theories. However, for the time being, we'll ignore genes and focus primarily on actual adaptations that evolved because they helped our ancestors survive and/or reproduce.

Evolutionary Psychology

Evolutionary psychology aims the lens of modern evolutionary theory on the workings of the human mind. It focuses primarily on **psychological adaptations**: mechanisms of the mind that have evolved to solve specific problems of survival or reproduction. These kinds of adaptations are in contrast to *physiological* adaptations, which are adaptations that occur in the body as a consequence of one's environment. One example of a physiological adaptation is how our skin makes calluses. First, there is an "input," such as repeated friction to the skin on the bottom of our feet from walking. Second, there is a "procedure," in which the skin grows new skin cells at the afflicted area. Third, an actual callus forms as an "output" to protect the underlying tissue—the final outcome of the physiological adaptation (i.e., tougher skin to protect repeatedly scraped areas). On the other hand, a *psychological* adaptation is a development or change of a mechanism in the mind. For example, take sexual jealousy. First, there is an "input," such as a romantic partner flirting with a rival. Second, there is a "procedure," in which the person evaluates the threat the rival poses to the romantic relationship. Third, there is a behavioral output, which might range from vigilance (e.g., snooping through a partner's email) to violence (e.g., threatening the rival). Although such behaviors serve a purpose for the jealous person, they can be harmful to others.

Evolutionary psychology is fundamentally an *interactionist* framework, or a theory that takes into account multiple factors when determining the outcome. For example, jealousy, like a callus, doesn't simply pop up out of nowhere. There is an "interaction" between the environmental trigger (e.g., the flirting; the repeated rubbing of the skin) and the initial response (e.g., evaluation of the flirter's threat; the forming of new skin cells) to produce the outcome.

In evolutionary psychology, culture also has a major effect on psychological adaptations. For example, status within one's group is important in all cultures for achieving reproductive success, because higher status makes someone more attractive to mates. In individualistic cultures, such as the United States, status is heavily determined by individual accomplishments. But in more collectivist cultures, such as Japan, status is more heavily determined by contributions to the group and by that group's success. For example, consider a group project. If you were to put in most of the effort on a successful group project, the culture in the United States reinforces the psychological adaptation to try to claim that success for yourself (because individual achievements are rewarded with higher status). However, the culture in Japan reinforces the psychological adaptation to attribute that success to the whole group (because collective achievements are rewarded with higher status). Another example of cultural input is the importance of virginity as a desirable quality for a mate. Cultural norms that advise against premarital sex persuade people to ignore their own basic interests because they know that virginity will make them more attractive marriage partners. Evolutionary psychology, in short, does not predict rigid robotic-like "instincts." That is, there isn't one rule that works all the time. Rather, evolutionary psychology studies flexible, environmentally-connected and culturally-influenced adaptations that vary according to the situation.

Psychological adaptations are hypothesized to be wide-ranging, and include food preferences, habitat preferences, mate preferences, and specialized fears. These psychological adaptations also include many traits that improve people's ability to live in groups, such as the desire to cooperate and make friends, or the inclination to spot and avoid frauds, punish rivals, establish status hierarchies, nurture children, and help genetic relatives. Research programs in evolutionary psychology develop and empirically test predictions about the nature of psychological adaptations. Below, we highlight a few evolutionary psychological theories and their associated research approaches.

Sexual Strategies Theory

Sexual strategies theory is based on sexual selection theory. It proposes that humans have evolved a list of different mating strategies, both short-term and long-term, that vary depending on culture, social context, parental influence, and personal mate value (desirability in the "mating market").

In its initial formulation, sexual strategies theory focused on the differences between men and women in mating preferences and strategies (Buss & Schmitt, 1993). It started by looking at the minimum parental investment needed to produce a child. For women, even the minimum investment is significant: after becoming pregnant, they have to carry that child for nine months inside of them. For men, on the other hand, the minimum investment to produce the same child is considerably smaller—simply the act of sex.

These differences in parental investment have an enormous impact on sexual strategies. For a woman, the risks associated with making a poor mating choice is high. She might get pregnant by a man who will not help to support her and her children, or who might have poor-quality genes. And because the stakes are higher for a woman, *wise* mating decisions for her are much more valuable. For men, on the other hand, the need to focus on making wise mating decisions isn't as important. That is, unlike women, men 1) don't biologically have the child growing inside of them for nine months, and 2) do not have as high a cultural expectation to raise the child. This logic leads to a powerful set of predictions: In short-term mating, women will likely be choosier than men (because the costs of getting pregnant are so high), while men, on average, will likely engage in more casual sexual activities (because this cost is greatly lessened). Due to this, men will sometimes deceive women about their long-term intentions for the benefit of short-term sex, and men are more likely than women to lower their mating standards for short-term mating situations.

Because women bear responsibility for pregnancy, they may use different sexual selection strategies than men do. [Image: CC0 Public Domain, https://goo.gl/m25gce]

An extensive body of empirical evidence supports these and related predictions (Buss & Schmitt, 2011). Men express a desire for a larger number of sex partners than women do. They let less time elapse before seeking sex. They are more willing to consent to sex with strangers and are less likely to require emotional involvement with their sex partners. They have more frequent sexual fantasies and fantasize about a larger variety of sex partners. They are more likely to regret missed sexual opportunities. And they lower their standards in short-term mating, showing a willingness to mate with a larger variety of women as long as the costs and risks are low.

However, in situations where both the man and woman are interested in long-term mating, both sexes tend to invest substantially in the relationship and in their children. In these cases, the theory predicts that both sexes will be extremely choosy when pursuing a long-term mating strategy. Much empirical research supports this prediction, as well. In fact, the qualities women and men generally look for when choosing long-term mates are very similar: both want mates who are intelligent, kind, understanding, healthy, dependable, honest, loyal, loving, and adaptable.

Nonetheless, women and men do differ in their preferences for a few key qualities in long-term mating, because of somewhat distinct adaptive problems. Modern women have inherited the evolutionary trait to desire mates who possess resources, have qualities linked with acquiring resources (e.g., ambition, wealth, industriousness), and are willing to share those resources with them. On the other hand, men more strongly desire youth and health in women, as both are cues to fertility. These male and female differences are universal in humans. They were first documented in 37 different cultures, from Australia to Zambia (Buss, 1989), and have been replicated by dozens of researchers in dozens of additional cultures (for summaries, see Buss, 2012).

As we know, though, just because we have these mating preferences (e.g., men with resources; fertile women), people don't always get what they want. There are countless other factors which influence who people ultimately select as their mate. For example, the sex ratio (the percentage of men to women in the mating pool), cultural practices (such as arranged marriages, which inhibit individuals' freedom to act on their preferred mating strategies), the strategies of others (e.g., if everyone else is pursuing short-term sex, it's more difficult to pursue a long-term mating strategy), and many others all influence who we select as our mates.

Sexual strategies theory—anchored in sexual selection theory— predicts specific similarities and differences in men and women's mating preferences and strategies. Whether we seek short-term or long-term relationships, many personality, social, cultural, and ecological factors will all influence who our partners will be.

Error Management Theory

Error management theory (EMT) deals with the evolution of how we think, make decisions, and evaluate uncertain situations—that is, situations where there's no clear answer how we should behave. (Haselton & Buss, 2000; Haselton, Nettle, & Andrews, 2005). Consider, for example, walking through the woods at dusk. You hear a rustle in the leaves on the path in front of you. It could be a snake. Or, it could just be the wind blowing the leaves. Because you can't really tell why the leaves rustled, it's an uncertain situation. The important question then is, what are the costs of errors in judgment? That is, if you conclude that it's a dangerous snake so you avoid the leaves, the costs are minimal (i.e., you simply make a short detour around them). However, if you assume the leaves are safe and simply walk over them—when in fact

If you were walking in the woods and heard a sound in the bushes you might be startled and act on the worst case scenario—such as the threat of a wild animal—by moving in the opposite direction. This is evolutionary psychology at work, keeping you safe so you can survive and reproduce. [Image: Nicholas T, https://goo.gl/gZ3zEL, CC BY 2.0, https://goo.gl/9uSnqN]

it *is* a dangerous snake—the decision could cost you your life.

Now, think about our evolutionary history and how generation after generation was confronted with similar decisions, where one option had low cost but great reward (walking around the leaves and not getting bitten) and the other had a low reward but high cost (walking through the leaves and getting bitten). These kinds of choices are called "cost asymmetries." If during our evolutionary history we encountered decisions like these generation after generation, over time an adaptive bias would be created: we would make sure to err in favor of the least costly (in this case, least dangerous) option (e.g., walking around the leaves). To put it another way, EMT predicts that whenever uncertain situations present us with a safer versus more dangerous decision, we will psychologically adapt to prefer choices that minimize the cost of errors.

EMT is a general evolutionary psychological theory that can be applied to many different domains of our lives, but a specific example of it is the *visual descent illusion*. To illustrate: Have you ever thought it would be no problem to jump off of a ledge, but as soon as you stood up there, it suddenly looked much higher than you thought? The visual descent illusion (Jackson & Cormack, 2008) states that people will overestimate the distance when looking down from a height (compared to looking up) so that people will be especially wary of falling from great heights—which would result in injury or death. Another example of EMT is the *auditory looming bias*: Have you ever noticed how an ambulance seems closer when it's coming toward you, but suddenly seems far away once it's immediately passed? With the auditory looming bias, people overestimate how close objects are when the sound is moving toward them compared to when it is moving away from them. From our evolutionary history, humans learned, "It's better to be safe than sorry." Therefore, if we think that a threat is closer to us when it's moving toward us (because it seems louder), we will be quicker to act and escape. In this regard, there may be times we ran away when we didn't need to (a false alarm), but wasting that time is a less costly mistake than not acting in the first place when a real threat does exist.

EMT has also been used to predict adaptive biases in the domain of mating. Consider something as simple as a smile. In one case, a smile from a potential mate could be a sign of sexual or romantic interest. On the other hand, it may just signal friendliness. Because of the costs to men of missing out on chances for reproduction, EMT predicts that men have a *sexual overperception bias*: they often misread sexual interest from a woman, when really it's just a friendly smile or touch. In the mating domain, the sexual overperception bias is one of the best-documented phenomena. It's been shown in studies in which men and women rated the sexual interest between people in photographs and videotaped interactions. As well, it's been shown in the laboratory with participants engaging in actual "speed dating," where the men interpret sexual interest from the women more often than the women actually intended it (Perilloux, Easton, & Buss, 2012). In short, EMT predicts that men, more than women, will over-infer sexual interest based on minimal cues, and empirical research confirms this adaptive mating bias.

Conclusion

Sexual strategies theory and error management theory are two evolutionary psychological theories that have received much empirical support from dozens of independent researchers. But, there are many other evolutionary psychological theories, such as social exchange theory for example, that also make predictions about our modern day behavior and preferences, too. The merits of each evolutionary psychological theory, however, must be evaluated separately and treated like any scientific theory. That is, we should only trust their predictions and claims to the extent they are supported by scientific studies. However, even if the theory is scientifically grounded, just because a psychological adaptation was advantageous in our history, it doesn't mean it's still useful today. For example, even though women may have preferred men with resources in generations ago, our modern society has advanced such that these preferences are no longer apt or necessary. Nonetheless, it's important to consider how our evolutionary history has shaped our automatic or "instinctual" desires and reflexes of today, so that we can better shape them for the future ahead.

Outside Resources

FAQs
http://www.anth.ucsb.edu/projects/human/evpsychfaq.html

Web: Articles and books on evolutionary psychology
http://homepage.psy.utexas.edu/homepage/Group/BussLAB/

Web: Main international scientific organization for the study of evolution and human behavior, HBES
http://www.hbes.com/

Discussion Questions

1. How does change take place over time in the living world?
2. Which two potential psychological adaptations to problems of survival are not discussed in this module?
3. What are the psychological and behavioral implications of the fact that women bear heavier costs to produce a child than men do?
4. Can you formulate a hypothesis about an error management bias in the domain of social interaction?

Vocabulary

Adaptations
Evolved solutions to problems that historically contributed to reproductive success.

Error management theory (EMT)
A theory of selection under conditions of uncertainty in which recurrent cost asymmetries of judgment or inference favor the evolution of adaptive cognitive biases that function to minimize the more costly errors.

Evolution
Change over time. Is the definition changing?

Gene Selection Theory
The modern theory of evolution by selection by which differential gene replication is the defining process of evolutionary change.

Intersexual selection
A process of sexual selection by which evolution (change) occurs as a consequences of the mate preferences of one sex exerting selection pressure on members of the opposite sex.

Intrasexual competition
A process of sexual selection by which members of one sex compete with each other, and the victors gain preferential mating access to members of the opposite sex.

Natural selection
Differential reproductive success as a consequence of differences in heritable attributes.

Psychological adaptations
Mechanisms of the mind that evolved to solve specific problems of survival or reproduction; conceptualized as information processing devices.

Sexual selection
The evolution of characteristics because of the mating advantage they give organisms.

Sexual strategies theory
A comprehensive evolutionary theory of human mating that defines the menu of mating strategies humans pursue (e.g., short-term casual sex, long-term committed mating), the

adaptive problems women and men face when pursuing these strategies, and the evolved solutions to these mating problems.

References

Buss, D. M. (2012). Evolutionary psychology: The new science of the mind (4th ed.). Boston, MA: Allyn & Bacon.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. Behavioral & Brain Sciences, 12, 1–49.

Buss, D. M., & Schmitt, D. P. (2011). Evolutionary psychology and feminism. Sex Roles, 64, 768–787.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. Psychological Review, 100, 204–232.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. Journal of Personality and Social Psychology, 78, 81–91.

Haselton, M. G., Nettle, D., & Andrews, P. W. (2005). The evolution of cognitive bias. In D. M. Buss (Ed.), The handbook of evolutionary psychology (pp. 724–746). New York, NY: Wiley.

Jackson, R. E., & Cormack, J. K. (2008). Evolved navigation theory and the environmental vertical illusion. Evolution and Human Behavior, 29, 299–304.

Perilloux, C., Easton, J. A., & Buss, D. M. (2012). The misperception of sexual interest. Psychological Science, 23, 146–151.

8
Hormones & Behavior

Randy J. Nelson

The goal of this module is to introduce you to the topic of hormones and behavior. This field of study is also called behavioral endocrinology, which is the scientific study of the interaction between hormones and behavior. This interaction is bidirectional: hormones can influence behavior, and behavior can sometimes influence hormone concentrations. Hormones are chemical messengers released from endocrine glands that travel through the blood system to influence the nervous system to regulate behaviors such as aggression, mating, and parenting of individuals.

Learning Objectives

- Define the basic terminology and basic principles of hormone–behavior interactions.
- Explain the role of hormones in behavioral sex differentiation.
- Explain the role of hormones in aggressive behavior.
- Explain the role of hormones in parental behavior.
- Provide examples of some common hormone–behavior interactions.

Introduction

This module describes the relationship between hormones and behavior. Many readers are likely already familiar with the general idea that hormones can affect behavior. Students are generally familiar with the idea that sex-hormone concentrations increase in the blood during puberty and decrease as we age, especially after about 50 years of age. Sexual behavior shows a similar pattern. Most people also know about the relationship between aggression and anabolic steroid hormones, and they know that administration of artificial steroid hormones sometimes results in uncontrollable, violent behavior called “roid rage.” Many different hormones can influence several types of behavior, but for the purpose of this module, we will restrict our discussion to just a few examples of hormones and behaviors. For example, are behavioral sex differences the result of hormones, the environment, or some combination of factors? Why are men much more likely than women to commit aggressive acts? Are hormones involved in mediating the so-called maternal “instinct”? Behavioral endocrinologists are interested in how the general physiological effects of hormones alter the development and expression of behavior and how behavior may influence the effects of hormones. This module describes, both phenomenologically and functionally, how hormones affect behavior.

To understand the hormone-behavior relationship, it is important briefly to describe hormones. **Hormones** are organic chemical messengers produced and released by specialized glands called **endocrine glands**. Hormones are released from these glands into the blood, where they may travel to act on target structures at some distance from their origin. Hormones are similar in function to **neurotransmitters**, the chemicals used by the nervous system in coordinating animals’ activities. However, hormones can operate over a greater distance and over a much greater temporal range than neurotransmitters (Focus Topic 1). Examples of hormones that influence behavior include steroid hormones such as **testosterone** (a common type of androgen), estradiol (a common type of estrogen), progesterone (a common type of **progestin**), and cortisol (a common type of glucocorticoid) (Table 1, A-B). Several types of protein or peptide (small protein) hormones also influence behavior, including **oxytocin**, vasopressin, **prolactin**, and leptin.

Focus Topic 1: Neural Transmission versus Hormonal Communication

Although neural and hormonal communication both rely on chemical signals, several prominent differences exist. Communication in the nervous system is analogous to traveling on a train. You can use the train in your travel plans as long as tracks exist between your proposed origin and destination. Likewise, neural messages can travel only to destinations along existing nerve tracts. Hormonal communication, on the other hand, is like traveling in a car. You can drive to many more destinations than train travel allows because there are many more roads than railroad tracks. Similarly, hormonal messages can travel anywhere in the body via the circulatory system; any cell receiving blood is potentially able to receive a hormonal message.

Neural and hormonal communication differ in other ways as well. To illustrate them, consider the differences between digital and analog technologies. Neural messages are digital, all-or-none events that have rapid onset and offset: neural signals can take place in milliseconds. Accordingly, the nervous system mediates changes in the body that are relatively rapid. For example, the nervous system regulates immediate food intake and directs body movement. In contrast, hormonal messages are analog, graded events that may take seconds, minutes, or even hours to occur. Hormones can mediate long-term processes, such as growth, development, reproduction, and metabolism.

Hormonal and neural messages are both chemical in nature, and they are released and received by cells in a similar manner; however, there are important differences as well. Neurotransmitters, the chemical messengers used by neurons, travel a distance of only 20–30 nanometers (30 X 10–9 m)—to the membrane of the postsynaptic neuron, where they bind with receptors. Hormones enter the circulatory system and may travel from 1 millimeter to >2 meters before arriving at a target cell, where they bind with specific receptors.

Another distinction between neural and hormonal communication is the degree of voluntary control that can be exerted over their functioning. In general, there is more voluntary control of neural than of hormonal signals. It is virtually impossible to will a change in your thyroid hormone levels, for example, whereas moving your limbs on command is easy.

Although these are significant differences, the division between the nervous system and the endocrine system is becoming more blurred as we learn more about how the nervous system regulates hormonal communication. A better understanding of the interface between the endocrine system and the nervous system, called neuroendocrinology, is likely to yield important advances in the future study of the interaction between hormones and behavior.

Steroid Hormones	
Cortisol	Increases carbohydrate metabolism; mediates stress responses
Estradiol	Uterine and other female tissue development; regulates sexual motivation and performance in females and males
Testosterone	Promotes sperm production and male secondary sexual characteristics; promotes sexual motivation and behavior, typically by being converted to estradiol

Table 1-A: Prominent Hormones That Influence Behavior

Peptides and Protein Hormones	
Oxytocin	Stimulates milk letdown and uterine contractions during birth; Promotes social bonding
Prolactin	Many actions relating to reproduction, water balance, and behavior associated with parental care
Thyroxine	Increases oxidation rates in tissue and affects neural development
Vasopressin	Increases water reabsorption in the kidney and affects learning and memory

Table 1-B: Prominent Hormones That Influence Behavior

Hormones coordinate the physiology and behavior of individuals by regulating, integrating, and controlling bodily functions. Over evolutionary time, hormones have often been co-opted by the nervous system to influence behavior to ensure reproductive success. For example, the same hormones, testosterone and estradiol, that cause gamete (egg or sperm) maturation also promote mating behavior. This dual hormonal function ensures that mating behavior occurs when animals have mature gametes available for fertilization. Another example of endocrine regulation of physiological and behavioral function is provided by pregnancy. Estrogens and progesterone concentrations are elevated during pregnancy, and these hormones are often involved in mediating **maternal behavior** in the mothers.

Not all cells are influenced by each and every hormone. Rather, any given hormone can directly influence only cells that have specific hormone **receptors** for that particular hormone. Cells that have these specific receptors are called **target cells** for the hormone. The interaction of a hormone with its receptor begins a series of cellular events that eventually lead to activation of enzymatic pathways or, alternatively, turns on or turns off gene activation that regulates protein synthesis. The newly synthesized proteins may activate or deactivate other genes, causing yet another cascade of cellular events. Importantly, sufficient numbers of appropriate hormone receptors must be available for a specific hormone to produce any effects. For example, testosterone is important for male sexual behavior. If men have too little testosterone, then sexual motivation may be low, and it can be restored by testosterone treatment. However, if men have normal or even elevated levels of testosterone yet display low sexual drive, then it might be possible for a lack of receptors to be the cause and treatment with additional hormones will not be effective.

How might hormones affect behavior? In terms of their behavior, one can think of humans and other animals conceptually as comprised of three interacting components: (1) input systems (sensory systems), (2) integrators (the central nervous system), and (3) output systems, or effectors (e.g., muscles). Hormones do not *cause* behavioral changes. Rather, hormones influence these three systems so that specific stimuli are more likely to elicit certain responses in the appropriate behavioral or social context. In other words, hormones change the probability that a particular behavior will be emitted in the appropriate situation (Nelson, 2011). This is a critical distinction that can affect how we think of hormone-behavior relationships.

We can apply this three-component behavioral scheme to a simple behavior, singing in zebra finches. Only male zebra finches sing. If the testes of adult male finches are removed, then the birds reduce singing, but castrated finches resume singing if the testes are reimplanted, or if the birds are treated with either testosterone or estradiol. Although we commonly consider androgens to be "male" hormones and estrogens to be "female" hormones, it is common for testosterone to be converted to estradiol in nerve cells (Figure 1). Thus, many male-like behaviors are associated with the actions of estrogens! Indeed, all estrogens must first be converted from androgens because of the typical biochemical synthesis process. If the converting enzyme is low or missing, then it is possible for females to produce excessive androgens and subsequently develop associated male traits. It is also possible for estrogens in the environment to affect the nervous system of animals, including people (e.g., Kidd et al., 2007). Again, singing behavior is most frequent when blood testosterone or estrogen concentrations are high. Males sing to attract mates or ward off potential competitors from

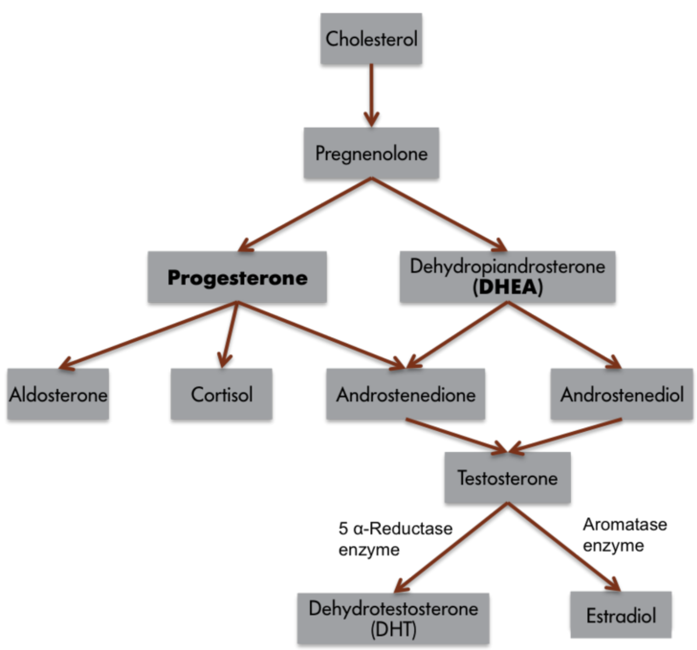
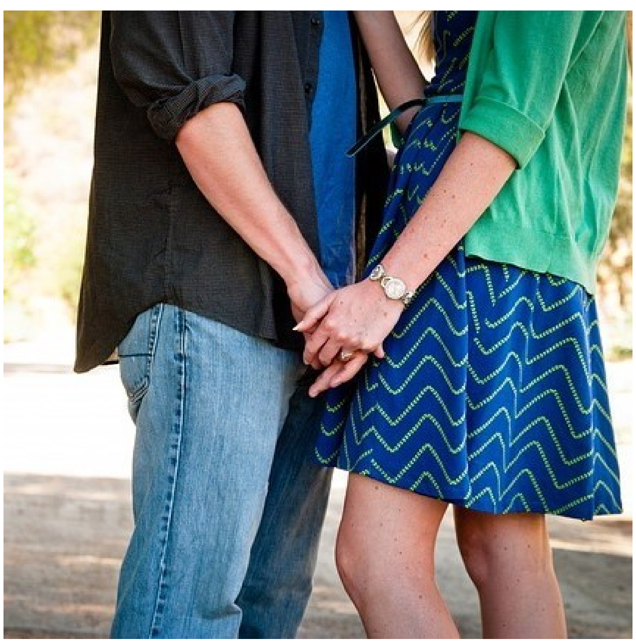


Figure 1: Biochemical Pathway for Steroid Hormone Synthesis: It is important to note that testosterone (an androgen) can be converted to another androgen, DHT, or an estrogen, estradiol. Too much or too little of the converting enzymes can influence brain and behavior.

their territories.

Although it is apparent from these observations that estrogens are somehow involved in singing, how might the three-component framework just introduced help us to formulate hypotheses to explore estrogen's role in this behavior? By examining input systems, we could determine whether estrogens alter the birds' sensory capabilities, making the environmental cues that normally elicit singing more salient. If this were the case, then females or competitors might be more easily seen or heard. Estrogens also could influence the central nervous system. Neuronal architecture or the speed of neural processing could change in the presence of estrogens. Higher neural processes (e.g., motivation, attention, or perception) also might be influenced. Finally, the effector organs, muscles in this case, could be affected by the presence of estrogens. Blood estrogen concentrations might somehow affect the muscles of a songbird's syrinx (the vocal organ of birds). Estrogens, therefore, could affect birdsong by influencing the sensory capabilities, central processing system, or effector organs of an individual bird. We do not understand completely how estrogen, derived from testosterone, influences birdsong, but in most cases, hormones can be considered to affect behavior by influencing one, two, or all three of these components, and this three-part framework can aid in the design of hypotheses and experiments to explore these issues.

How might behaviors affect hormones? The birdsong example demonstrates how hormones can affect behavior, but as noted, the reciprocal relation also occurs; that is, behavior can affect hormone concentrations. For example, the sight of a territorial intruder may elevate blood testosterone concentrations in resident male birds and thereby stimulate singing or fighting behavior. Similarly, male mice or rhesus monkeys that lose a fight decrease circulating testosterone concentrations for several days or even weeks afterward. Comparable results have also been reported in humans. Testosterone concentrations are affected not only in humans involved in physical combat, but also in those involved in simulated battles. For example, testosterone concentrations were elevated in winners and reduced in losers of regional chess tournaments.



The expectation of events can influence one's hormonal activity. How do you think yours is affected if you anticipate going on a date with a romantic interest soon? [Image: CC0 Public Domain, https://goo.gl/m25gce]

People do not have to be directly involved in a contest to have their hormones affected by the outcome of the contest. Male fans of both the Brazilian and Italian teams were recruited to provide saliva samples to be assayed for testosterone before and after the final game of the World Cup soccer match in 1994. Brazil and Italy were tied going into the final game, but Brazil won on a penalty kick at the last possible moment. The Brazilian fans were elated and the Italian fans were crestfallen. When the samples were assayed, 11 of 12 Brazilian fans who were sampled had increased testosterone concentrations, and 9 of 9 Italian fans had decreased testosterone concentrations, compared with pre-game baseline values (Dabbs, 2000).

In some cases, hormones can be affected by anticipation of behavior. For example, testosterone concentrations also influence sexual motivation and behavior in women. In one study, the interaction between sexual intercourse and testosterone was compared with other activities (cuddling or exercise) in women (van Anders, Hamilton, Schmidt, & Watson, 2007). On three separate occasions, women provided a pre-activity, post-activity, and next-morning saliva sample. After analysis, the women's testosterone was determined to be elevated prior to intercourse as compared to other times. Thus, an anticipatory relationship exists between sexual behavior and testosterone. Testosterone values were higher post-intercourse compared to exercise, suggesting that engaging in sexual behavior may also influence hormone concentrations in women.

Sex Differences

Hens and roosters are different. Cows and bulls are different. Men and women are different. Even girls and boys are different. Humans, like many animals, are sexually dimorphic (*di*, "two"; *morph*, "type") in the size and shape of their bodies, their physiology, and for our purposes, their behavior. The behavior of boys and girls differs in many ways. Girls generally excel in verbal abilities relative to boys; boys are nearly twice as likely as girls to suffer from dyslexia (reading difficulties) and stuttering and nearly 4 times more likely to suffer from autism. Boys are generally better than girls at tasks that require visuospatial abilities. Girls engage in nurturing behaviors more frequently than boys. More than 90% of all anorexia nervosa cases involve young women. Young men are twice as likely as young women to suffer from schizophrenia. Boys are much more aggressive and generally engage in more rough-and-tumble play than girls (Berenbaum, Martin, Hanish, Briggs, & Fabes, 2008). Many sex differences, such as the difference in aggressiveness, persist throughout adulthood. For example, there are many more men than women serving prison sentences for violent behavior. The hormonal differences between men and women may account for adult sex differences that develop during puberty, but what accounts for behavioral sex differences among children *prior* to puberty and activation of their gonads? Hormonal secretions from the developing gonads determine whether the individual develops in a male or female manner. The mammalian embryonic testes produce androgens, as well as peptide hormones, that steer the development of the body, central nervous system, and subsequent behavior in a male direction. The embryonic ovaries of mammals are virtually quiescent and do not secrete high concentrations of hormones. In the presence of ovaries, or in the complete absence of any gonads, morphological, neural, and, later, behavioral development follows a female pathway.

Gonadal steroid hormones have organizational (or programming) effects upon brain and behavior (Phoenix, Goy, Gerall, & Young, 1959). The organizing effects of steroid hormones are relatively constrained to the early stages of development. An asymmetry exists in the effects of testes and ovaries on the organization of behavior in mammals. Hormone exposure early in life has organizational effects on subsequent rodent behavior; early steroid hormone treatment causes relatively irreversible and permanent **masculinization** of rodent behavior (mating and aggressive). These early hormone effects can be contrasted with the reversible behavioral influences of steroid hormones provided in adulthood, which are called activational effects. The activational effects of hormones on adult behavior are temporary and may wane soon after the hormone is metabolized. Thus, typical male behavior requires exposure to androgens during gestation (in humans) or immediately after birth (in rodents) to somewhat masculinize the brain and also requires androgens during or after puberty to activate these neural circuits. Typical female behavior requires a lack of exposure to androgens early in life which leads to **feminization** of the brain and also requires estrogens to activate these neural circuits in adulthood. But this simple dichotomy, which works well with animals with very distinct sexual dimorphism in behavior, has many caveats when applied to people.

Sex differences in appearance are often more pronounced in nonhuman animals than in humans. Male birds particularly, for example roosters, tend to have physical features that differ from the females and also differ significantly in size. [Image: John Cudworth, https://goo.gl/oopnqM, CC BY-NC 2.0, https://goo.gl/VnKlK8]

If you walk through any major toy store, then you will likely observe a couple of aisles filled with pink boxes and the complete absence of pink packaging of toys in adjacent aisles. Remarkably, you will also see a strong self-segregation of boys and girls in these aisles. It is rare to see boys in the "pink" aisles and vice versa. The toy manufacturers are often accused of making toys that are gender biased, but it seems more likely that boys and girls enjoy playing with specific types and colors of toys. Indeed, toy manufacturers would immediately double their sales if they could sell toys to both sexes. Boys generally prefer toys such as trucks and balls and girls generally prefer toys such as dolls. Although it is doubtful that there are genes that encode preferences for toy cars and trucks on the Y chromosome, it is possible that hormones might shape the development of a child's brain to prefer certain types of toys or styles of play behavior. It is reasonable to believe that children learn which types of toys and which styles of play are appropriate to their gender. How can we understand and separate the contribution of physiological mechanisms from learning to understand sex differences in human behaviors? To untangle these issues, animal models are often used. Unlike the situation in humans, where sex differences are usually only a matter of degree (often slight), in some animals, members of only one sex may display a particular behavior. As noted, often only male songbirds sing. Studies of such strongly sex-biased behaviors are particularly valuable for understanding the interaction among behavior, hormones, and the nervous system.

A study of vervet monkeys calls into question the primacy of learning in the establishment of toy preferences (Alexander & Hines, 2002). Female vervet monkeys preferred girl-typical toys, such as dolls or cooking pots, whereas male vervet monkeys preferred boy-typical toys, such as cars or balls. There were no sex differences in preference for gender-neutral toys, such as picture books or stuffed animals. Presumably, monkeys have no prior concept of "boy" or "girl" toys. Young rhesus monkeys also show similar toy preferences.

If you think back to the toys and clothing you played with and wore in your youth, do you think they were more a result of your hormonal activity or the choices that society and your parents made for you? [Image: CC0 Public Domain, https://goo.gl/m25gce]

What then underlies the sex difference in toy preference? It is possible that certain attributes of toys (or objects) appeal to either boys or girls. Toys that appeal to boys or male vervet or rhesus monkeys, in this case, a ball or toy car, are objects that can be moved actively through space, toys that can be incorporated into active, rough and tumble play. The appeal of toys that girls or female vervet monkeys prefer appears to be based on color. Pink and red (the colors of the doll and pot) may provoke attention to infants.

Society may reinforce such stereotypical responses to gender-typical toys. The sex differences in toy preferences emerge by 12 or 24 months of age and seem fixed by 36 months of age, but are sex differences in toy preference present during the first year of life? It is difficult to ask pre-verbal infants what they prefer, but in studies where the investigators examined the amount of time that babies looked at different toys, eye-tracking data indicate that infants as young as 3 months showed sex differences in toy preferences; girls preferred dolls, whereas boys preferred trucks. Another result that suggests, but does not prove, that hormones are involved in toy preferences is the observation that girls diagnosed with congenital adrenal hyperplasia (CAH), whose adrenal glands produce varying amounts of androgens early in life, played with masculine toys more often than girls without CAH. Further, a dose-response relationship between the extent of the disorder (i.e., degree of fetal androgen exposure) and degree of masculinization of play behavior was observed. Are the sex differences in toy preferences or play activity, for example, the inevitable consequences of the differential endocrine environments of boys and girls, or are these differences imposed by cultural practices and beliefs? Are these differences the result of receiving gender-specific toys from an early age, or are these differences some combination of endocrine and cultural factors? Again, these are difficult questions to unravel in people.

Even when behavioral sex differences appear early in development, there seems to be some question regarding the influences of societal expectations. One example is the pattern of human play behavior during which males are more physical; this pattern is seen in a number of other species including nonhuman primates, rats, and dogs. Is the difference in the frequency of rough-and-tumble play between boys and girls due to biological factors associated with being male or female, or is it due to cultural expectations and learning? If there is a combination of biological and cultural influences mediating the frequency of rough-and-tumble play, then what proportion of the variation between the sexes is due to biological factors and what proportion is due to social influences? Importantly, is it appropriate to talk about “normal” sex differences when these traits virtually always arrange themselves along a continuum rather than in discrete categories?

Sex differences are common in humans and in nonhuman animals. Because males and females differ in the ratio of androgenic and estrogenic steroid hormone concentrations, behavioral endocrinologists have been particularly interested in the extent to which behavioral sex differences are mediated by hormones. The process of becoming female or male is called **sexual differentiation**. The primary step in sexual differentiation occurs at fertilization. In mammals, the ovum (which always contains an X chromosome) can be fertilized by a sperm bearing either a Y or an X chromosome; this process is called **sex determination**. The **chromosomal sex** of homogametic mammals (XX) is female; the chromosomal sex of heterogametic mammals (XY) is male. Chromosomal sex determines **gonadal sex**. Virtually all subsequent sexual differentiation is typically the result of differential exposure to gonadal steroid hormones. Thus, gonadal sex determines hormonal sex, which regulates morphological sex. Morphological differences in the central nervous system, as well as in some effector organs, such as muscles, lead to behavioral sex differences. The process of sexual differentiation is complicated, and the potential for errors is present. Perinatal exposure to androgens is the most common cause of anomalous sexual differentiation among females. The source of androgen may be internal (e.g., secreted by the adrenal glands) or external (e.g., exposure to environmental estrogens). Turner syndrome results when the second X chromosome is missing or damaged; these individuals possess dysgenic ovaries and are not exposed to steroid hormones until puberty. Interestingly, women with Turner syndrome often have impaired spatial memory.

Female mammals are considered the "neutral" sex; additional physiological steps are required for male differentiation, and more steps bring more possibilities for errors in differentiation. Some examples of male anomalous sexual differentiation include **5α-reductase** deficiency (in which XY individuals are born with ambiguous genitalia because of a lack of dihydrotestosterone and are reared as females, but masculinization occurs during puberty) and androgen insensitivity syndrome or TFM (in which XY individuals lack receptors for androgens and develop as females). By studying individuals who do not neatly fall into the dichotic boxes of female or male and for whom the process of sexual differentiation is atypical, behavioral endocrinologists glean hints about the process of typical sexual differentiation.

We may ultimately want to know how hormones mediate sex differences in the human brain and behavior (to the extent to which these differences occur). To understand the mechanisms underlying sex differences in the brain and behavior, we return to the birdsong example. Birds provide the best evidence that behavioral sex differences are the result of hormonally induced structural changes in the brain (Goodson, Saldanha, Hahn, & Soma, 2005). In contrast to mammals, in which structural differences in neural tissues have not been directly linked to behavior, structural differences in avian brains have been directly linked to a sexually behavior: birdsong.

Several brain regions in songbirds display significant sex differences in size. Two major brain circuit pathways, (1) the song production motor pathway and (2) the auditory transmission pathway, have been implicated in the learning and production of birdsong. Some parts of the song production pathway of male zebra finches are 3 to 6 times larger than those of female conspecifics. The larger size of these brain areas reflects that neurons in these nuclei are larger, more numerous, and farther apart. Although castration of adult male birds reduces singing, it does not reduce the size of the brain nuclei controlling song production. Similarly, androgen treatment of adult female zebra finches does not induce changes either in singing or in the size of the song control regions. Thus, activational effects of steroid hormones do not account for the sex differences in singing behavior or brain nucleus size in zebra finches. The sex differences in these structures are organized or programmed in the egg by estradiol (masculinizes) or the lack of steroids (feminizes).

Taken together, estrogens appear to be necessary to activate the neural machinery underlying the song system in birds. The testes of birds primarily produce androgens, which enter the circulation. The androgens enter neurons containing aromatase, which converts them to estrogens. Indeed, the brain is the primary source of estrogens, which activate masculine behaviors in many bird species.

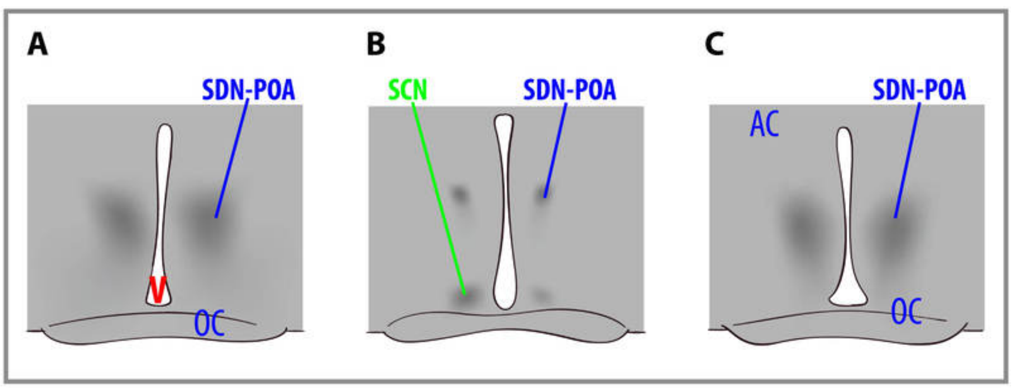


Figure 2: The sexually dimorphic nuclei of the preoptic area (SDN-POA) Gonadal steroid hormones have organizing effects upon brain and behavior. The organizing effects of steroid hormones are relatively constrained to the early stages of development. Exposure to testosterone (which is converted to estradiol) or estradiol causes masculinization of the brain. These are cross-sections through the brains of rats that show a male (left), a female (center), and a female treated with testosterone as a newborn (right). Note that the SDN-POA (the dark cell bodies) of the male are substantially larger than those of the untreated female but are equal in size to those of the testosterone-treated female. The extent that these sex differences in brain structure account for sex differences in behavior remain unspecified in mammals. OC = optic chiasm; SCN = suprachiasmatic nucleus; V = third ventricle.

Sex differences in human brain size have been reported for years. More recently, sex differences in specific brain structures have been discovered (Figure 2). Sex differences in a number of cognitive functions have also been reported. Females are generally more sensitive to auditory information, whereas males are more sensitive to visual information. Females are also typically more sensitive than males to taste and olfactory input. Women display less lateralization of cognitive functions than men. On average, females generally excel in verbal, perceptual, and fine motor skills, whereas males outperform females on quantitative and visuospatial tasks, including map reading and direction finding. Although reliable sex differences can be documented, these differences in ability are slight. It is important to note that there is more variation *within* each sex than *between* the sexes for most cognitive abilities (Figure 3).

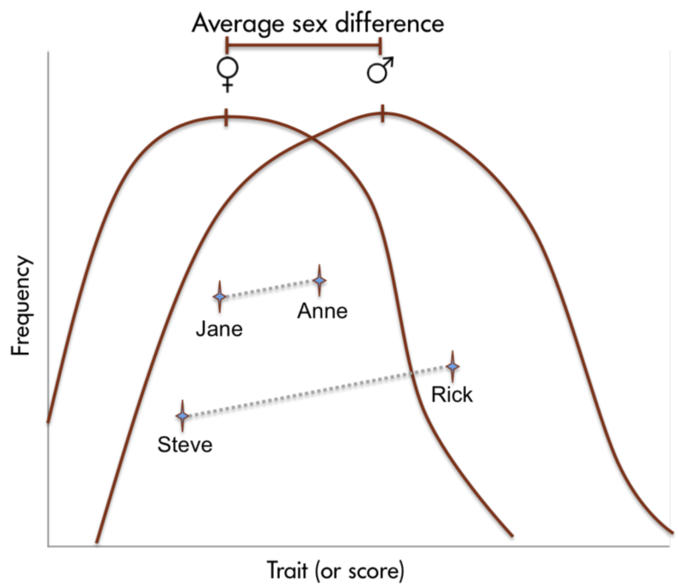


Figure 3: The average sex differences in human performance often reflect significant overlap between the sexes There are often greater differences in performance between individuals of the same sex (for example, between Steve and Rick in the figure) than between individuals of the opposite sex (for example, between Steve and Jane in the figure).

Aggressive Behaviors

The possibility for aggressive behavior exists whenever the interests of two or more individuals are in conflict (Nelson, 2006). Conflicts are most likely to arise over limited resources such as territories, food, and mates. A social interaction decides which animal gains access to the contested resource. In many cases, a submissive posture or gesture on the part of one animal avoids the necessity of actual combat over a resource. Animals may also participate in threat displays or ritualized combat in which dominance is determined but no physical damage is inflicted.

There is overwhelming circumstantial evidence that androgenic steroid hormones mediate aggressive behavior across many species. First, seasonal variations in blood plasma concentrations of testosterone and seasonal variations in aggression coincide. For instance, the incidence of aggressive behavior peaks for male deer in autumn, when they are secreting high levels of testosterone. Second, aggressive behaviors increase at the time of puberty, when the testes become active and blood concentrations of androgens rise. Juvenile deer do not participate in the fighting during the mating season. Third, in any given species, males are generally more aggressive than females. This is certainly true of deer; relative to stags, female deer rarely display aggressive behavior, and their rare aggressive acts are qualitatively different from the aggressive behavior of aggressive males. Finally, castration typically reduces aggression in males, and testosterone replacement therapy restores aggression to pre-castration levels. There are some interesting exceptions to these general observations that are outside the scope of this module.

As mentioned, males are generally more aggressive than females. Certainly, human males are much more aggressive than females. Many more men than women are convicted of violent crimes in North America. The sex differences in human aggressiveness appear very early. At every age throughout the school years, many more boys than girls initiate physical assaults. Almost everyone will acknowledge the existence of this sex difference, but assigning a cause to behavioral sex differences in humans always elicits much debate. It is possible that boys are more aggressive than girls because androgens promote aggressive behavior and boys have higher blood concentrations of androgens than girls. It is possible that boys and girls differ in their aggressiveness because the brains of boys are exposed to androgens prenatally and the “wiring” of their brains is thus organized in a way that facilitates the expression of aggression. It is also possible that boys are encouraged and girls are discouraged by family, peers, or others from acting in an aggressive manner. These three hypotheses are not mutually exclusive, but it is extremely difficult to discriminate among them to account for sex differences in human aggressiveness.

Researchers have electrically stimulated particular regions in people’s brains, and these individuals have burst into aggressive, violent behavior, helping demonstrate that such responses are hardwired into us. [Image: Riccardo Cuppini, https://goo.gl/b6bHU2, CC BY-NC-ND 2.0, https://goo.gl/bhtmIY]

What kinds of studies would be necessary to assess these hypotheses? It is usually difficult to separate out the influences of environment and physiology on the development of behavior in humans. For example, boys and girls differ in their rough-and-tumble play at a very young age, which suggests an early physiological influence on aggression. However, parents interact with their male and female offspring differently; they usually play more roughly with male infants than with females, which suggests that the sex difference in aggressiveness is partially learned. This difference in parental interaction style is evident by the first week of life. Because of these complexities in the factors influencing human behavior, the study of hormonal effects on sex-differentiated behavior has been pursued in nonhuman animals, for which environmental influences can be held relatively constant. Animal models for which sexual differentiation occurs postnatally are often used so that this process can be easily manipulated experimentally.

Again, with the appropriate animal model, we can address the questions posed above: Is the sex difference in aggression due to higher adult blood concentrations of androgens in males than in females, or are males more aggressive than females because their brains are organized differently by perinatal hormones? Are males usually more aggressive than females because of an interaction of early and current blood androgen concentrations? If male mice are castrated prior to their sixth day of life, then treated with testosterone propionate in adulthood, they show low levels of aggression. Similarly, females ovariectomized prior to their sixth day but given androgens in adulthood do not express male-like levels of aggression. Treatment of perinatally gonadectomized males or females with testosterone prior to their sixth day life and also in adulthood results in a level of aggression similar to that observed in typical male mice. Thus, in mice, the proclivity for males to act more aggressively than females is organized perinatally by androgens but also requires the presence of androgens after puberty in order to be fully expressed. In other words, aggression in male mice is both organized and activated by androgens. Testosterone exposure in adulthood without prior organization of the brain by steroid hormones does not evoke typical male levels of aggression. The hormonal control of aggressive behavior in house mice is thus similar to the hormonal mediation of heterosexual male mating behavior in other rodent species. Aggressive behavior is both organized and activated by androgens in many species, including rats, hamsters, voles, dogs, and possibly some primate species.

Parental Behaviors

Parental behavior can be considered to be any behavior that contributes directly to the survival of fertilized eggs or offspring that have left the body of the female. There are many patterns of mammalian parental care. The developmental status of the newborn is an important factor driving the type and quality of parental care in a species. Maternal care is much more common than **paternal** care. The vast majority of research on the hormonal correlates of mammalian parental behavior has been conducted on rats. Rats bear altricial young, and mothers perform a cluster of stereotyped maternal behaviors, including nest building, crouching over the pups to allow nursing and to provide warmth, pup retrieval, and increased aggression directed at intruders. If you expose nonpregnant female rats (or males) to pups, their most common reaction is to huddle far away from them. Rats avoid new things (neophobia). However, if you expose adult rats to pups every day, they soon begin to behave maternally. This process is called concaveation or sensitization and it appears to serve to reduce the adult rats' fear of pups.

Of course a new mother needs to act maternal as soon as her offspring arrive—not in a week. The onset of maternal behavior in rats is mediated by hormones. Several methods of study, such as hormone removal and replacement therapy, have been used to determine the hormonal correlates of rat maternal behavior. A fast decline of blood concentrations of **progesterone** in late pregnancy after sustained high concentrations of this hormone, in combination with high concentrations of estradiol and probably prolactin and oxytocin, induces female rats to behave maternally almost immediately in the presence of pups. This pattern of hormones at parturition overrides the usual fear response of adult rats toward pups, and it permits the onset of maternal behavior. Thus, the so-called maternal "instinct" requires hormones to increase the approach tendency and lower the avoidance tendency. Laboratory strains of mice and rats are usually docile, but mothers can be quite aggressive toward animals that venture too close to their litter. Progesterone appears to be the primary hormone that induces this maternal aggression in rodents, but species differences exist. The role of maternal aggression in women's behavior has not been adequately described or tested.

Although cortisol may not directly increase maternal behaviors, the next time your mom gives you a hug you know one hormone to thank. [Image: Maria Grazia Montagnari, https://goo.gl/LY1Tq0, CC BY 2.0, https://goo.gl/BRvSA7]

A series of elegant experiments by Alison Fleming and her collaborators studied the endocrine correlates of the behavior of human mothers as well as the endocrine correlates of maternal attitudes as expressed in self-report questionnaires. Responses such as patting, cuddling, or kissing the baby were called affectionate behaviors; talking, singing, or cooing to the baby were considered vocal behaviors. Both affectionate and vocal behaviors were considered approach behaviors. Basic caregiving activities, such as changing diapers and burping the infants, were also recorded. In these studies, no relationship between hormone concentrations and maternal responsiveness, as measured by attitude questionnaires, was found. For example, most women showed an increasing positive self-image during early pregnancy that dipped during the second half of pregnancy, but recovered after parturition. A related dip in feelings of maternal engagement occurred during late pregnancy, but rebounded substantially after birth in most women. However, when behavior, rather than questionnaire responses, was compared with hormone concentrations, a different story emerged. Blood plasma concentrations of cortisol were positively associated with approach behaviors. In other words, women who had high concentrations of blood cortisol, in samples obtained immediately before or after nursing, engaged in more physically affectionate behaviors and talked more often to their babies than mothers with low cortisol concentrations. Additional analyses from this study revealed that the correlation was even greater for mothers that had reported positive maternal regard (feelings and attitudes) during gestation. Indeed, nearly half of the variation in maternal behavior among women could be accounted for by cortisol concentrations and positive maternal attitudes during pregnancy.

Presumably, cortisol does not induce maternal behaviors directly, but it may act indirectly on the quality of maternal care by evoking an increase in the mother's general level of arousal, thus increasing her responsiveness to infant-generated cues. New mothers with high cortisol concentrations were also more attracted to their infant's odors, were superior in identifying their infants, and generally found cues from infants highly appealing (Fleming, Steiner, & Corter, 1997).

The medial preoptic area is critical for the expression of rat maternal behavior. The amygdala appears to tonically inhibit the expression of maternal behavior. Adult rats are fearful of pups, a response that is apparently mediated by chemosensory information. Lesions of the amygdala or afferent sensory pathways from the vomeronasal organ to the amygdala disinhibit the expression of maternal behavior. Hormones or sensitization likely act to disinhibit the amygdala, thus permitting the occurrence of maternal behavior. Although correlations have been established, direct evidence of brain structural changes in human mothers remains unspecified (Fleming & Gonzalez, 2009).

Considered together, there are many examples of hormones influencing behavior and of behavior feeding back to influence hormone secretion. More and more examples of hormone–behavior interactions are discovered, including hormones in the mediation of food and fluid intake, social interactions, salt balance, learning and memory, stress coping, as well as psychopathology including depression, anxiety disorders, eating disorders, postpartum depression, and seasonal depression. Additional research should reveal how these hormone–behavior interactions are mediated.

Outside Resources

Book: Adkins-Regan, E. (2005). Hormones and animal social behavior. Princeton, NJ: Princeton University Press.

Book: Beach, F. A. (1948). Hormones and behavior. New York: Paul Hoeber.

Book: Beach, F. A. (1975). Behavioral endocrinology: An emerging discipline. American Scientist, 63: 178–187.

Book: Nelson, R. J. (2011). An introduction to behavioral endocrinology (4th ed.). Sunderland, MA: Sinauer Associates.

Book: Pfaff, D. W. (2009). Hormones, brain, and behavior (2nd ed.). New York: Academic Press.

Book: Pfaff, D. W., Phillips, I. M., & Rubin, R. T. (2005). Principles of hormone/behavior relations. New York: Academic Press.

Video: Endocrinology Video (Playlist) - This YouTube playlist contains many helpful videos on the biology of hormones, including reproduction and behavior. This would be a helpful resource for students struggling with hormone synthesis, reproduction, regulation of biological functions, and signaling pathways.
https://www.youtube.com/playlist?list=PLqTetbgey0aemiTfD8QkMsSUq8hQzv-vA

Video: Paul Zak: Trust, morality - and oxytocin- This Ted talk explores the roles of oxytocin in the body. Paul Zak discusses biological functions of oxytocin, like lactation, as well as potential behavioral functions, like empathy.
https://www.youtube.com/watch?v=rFAdlU2ETjU

Video: Sex Differentiation- This video discusses gonadal differentiation, including the role of androgens in the development of male features.
https://www.youtube.com/watch?v=ciQjo7bj-uQ

Video: The Teenage Brain Explained- This is a great video explaining the roles of hormones during puberty.
https://www.youtube.com/watch?v=hiduiTq1ei8

Web: Society for Behavioral Neuroendocrinology - This website contains resources on

current news and research in the field of neuroendocrinology.
http://sbn.org/home.aspx

Discussion Questions

1. What are some of the problems associated with attempting to determine causation in a hormone–behavior interaction? What are the best ways to address these problems?
2. Hormones cause changes in the rates of cellular processes or in cellular morphology. What are some ways that these hormonally induced cellular changes might theoretically produce profound changes in behavior?
3. List and describe some behavioral sex differences that you have noticed between boys and girls. What causes girls and boys to choose different toys? Do you think that the sex differences you have noted arise from biological causes or are learned? How would you go about establishing your opinions as fact?
4. Why is it inappropriate to refer to androgens as "male" hormones and estrogens as "female" hormones?
5. Imagine that you discovered that the brains of architects were different from those of non-architects—specifically, that the "drawstraightem nuclei" of the right temporal lobe were enlarged in architects as compared with non-architects. Would you argue that architects were destined to be architects because of their brain organization or that experience as an architect changed their brains? How would you resolve this issue?

Vocabulary

5α-reductase
An enzyme required to convert testosterone to 5α-dihydrotestosterone.

Aggression
A form of social interaction that includes threat, attack, and fighting.

Aromatase
An enzyme that converts androgens into estrogens.

Chromosomal sex
The sex of an individual as determined by the sex chromosomes (typically XX or XY) received at the time of fertilization.

Defeminization
The removal of the potential for female traits.

Demasculinization
The removal of the potential for male traits.

Dihydrotestosterone (DHT)
A primary androgen that is an androgenic steroid product of testosterone and binds strongly to androgen receptors.

Endocrine gland
A ductless gland from which hormones are released into the blood system in response to specific biological signals.

Estrogen
Any of the C18 class of steroid hormones, so named because of the estrus-generating properties in females. Biologically important estrogens include estradiol and estriol.

Feminization
The induction of female traits.

Gonadal sex
The sex of an individual as determined by the possession of either ovaries or testes. Females

have ovaries, whereas males have testes.

Hormone
An organic chemical messenger released from endocrine cells that travels through the blood to interact with target cells at some distance to cause a biological response.

Masculinization
The induction of male traits.

Maternal behavior
Parental behavior performed by the mother or other female.

Neurotransmitter
A chemical messenger that travels between neurons to provide communication. Some neurotransmitters, such as norepinephrine, can leak into the blood system and act as hormones.

Oxytocin
A peptide hormone secreted by the pituitary gland to trigger lactation, as well as social bonding.

Parental behavior
Behaviors performed in relation to one's offspring that contributes directly to the survival of those offspring

Paternal behavior
Parental behavior performed by the father or other male.

Progesterone
A primary progestin that is involved in pregnancy and mating behaviors.

Progestin
A class of C21 steroid hormones named for their progestational (pregnancy-supporting) effects. Progesterone is a common progestin.

Prohormone
A molecule that can act as a hormone itself or be converted into another hormone with different properties. For example, testosterone can serve as a hormone or as a prohormone for either dihydrotestosterone or estradiol.

Prolactin
A protein hormone that is highly conserved throughout the animal kingdom. It has many biological functions associated with reproduction and synergistic actions with steroid hormones.

Receptor
A chemical structure on the cell surface or inside of a cell that has an affinity for a specific chemical configuration of a hormone, neurotransmitter, or other compound.

Sex determination
The point at which an individual begins to develop as either a male or a female. In animals that have sex chromosomes, this occurs at fertilization. Females are XX and males are XY. All eggs bear X chromosomes, whereas sperm can either bear X or Y chromosomes. Thus, it is the males that determine the sex of the offspring.

Sex differentiation
The process by which individuals develop the characteristics associated with being male or female. Differential exposure to gonadal steroids during early development causes sexual differentiation of several structures including the brain.

Target cell
A cell that has receptors for a specific chemical messenger (hormone or neurotransmitter).

Testosterone
The primary androgen secreted by the testes of most vertebrate animals, including men.

References

Alexander, G. M. & Hines, M. (2002). Sex differences in response to children's toys in nonhuman primates (*Cercopithecus aethiops sabaeus*). *Evolution and Human Behavior*, 23, 467–479.

Berenbaum, S. A., Martin, C. L., Hanish, L. D., Briggs, P. T., & Fabes, R. A. (2008). Sex differences in children's play. In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, J. Herman, & E. Young (Eds.), *Sex differences in the brain: From genes to behavior*. New York: Oxford University Press.

Dabbs, J. M. (2000). *Heroes, rogues, and lovers: Testosterone and behavior*. Columbus, OH: McGraw Hill.

Fleming, A. S., & Gonzalez, A. (2009). Neurobiology of human maternal care. In P. T. Ellison & P. B. Gray (Eds.), *Endocrinology of social relationships* (pp. 294–318). Cambridge, MA: Harvard University Press.

Fleming, A. S., Steiner, M., & Corter, C. (1997). Cortisol, hedonics, and maternal responsiveness in human mothers. *Hormones and Behavior*, 32, 85–98.

Goodson, J. L., Saldanha, C. J., Hahn, T. P., Soma, K. K. (2005). Recent advances in behavioral neuroendocrinology: Insights from studies on birds. *Hormones and Behavior*, 48, 461–73.

Kidd, K. A., Blanchfield, P. J., Mills, K. H., Palace, V. P., Evans, R. E. Lazorchak, J. M. & Flick, R. (2007). Collapse of a fish population following exposure to a synthetic estrogen. Proceedings of the National Academy of Sciences,104, 8897–8901.

Nelson, R. J. (Ed.) (2006). *Biology of aggression*. New York: Oxford University Press.

Nelson, R.J. (2011). *An introduction to behavioral endocrinology* (4th ed.). Sunderland, MA: Sinauer Associates.

Phoenix, C. H., Goy, R. W., Gerall, A. A., & Young, W. C. (1959). Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. *Endocrinology*, 65: 369–382.

van Anders, S., Hamilton, L., Schmidt, N., & Watson, N. (2007). Associations between testosterone secretion and sexual activity in women. *Hormones and Behavior*, 51, 477–82.

9
Biochemistry of Love

Sue Carter & Stephen Porges

Love is deeply biological. It pervades every aspect of our lives and has inspired countless works of art. Love also has a profound effect on our mental and physical state. A “broken heart” or a failed relationship can have disastrous effects; bereavement disrupts human physiology and may even precipitate death. Without loving relationships, humans fail to flourish, even if all of their other basic needs are met. As such, love is clearly not “just” an emotion; it is a biological process that is both dynamic and bidirectional in several dimensions. Social interactions between individuals, for example, trigger cognitive and physiological processes that influence emotional and mental states. In turn, these changes influence future social interactions. Similarly, the maintenance of loving relationships requires constant feedback through sensory and cognitive systems; the body seeks love and responds constantly to interactions with loved ones or to the absence of such interactions. The evolutionary principles and ancient hormonal and neural systems that support the beneficial and healing effects of loving relationships are described here.

Learning Objectives

- Understand the role of Oxytocin in social behaviors.
- Articulate the functional differences between Vasopressin and Oxytocin.
- List sex differences in reaction to stress.

Introduction

Although evidence exists for the healing power of love, only recently has science turned its attention to providing a physiological explanation for love. The study of love in this context offers insight into many important topics, including the biological basis of interpersonal relationships and why and how disruptions in social bonds have such pervasive consequences for behavior and physiology. Some of the answers will be found in our growing knowledge of the neurobiological and endocrinological mechanisms of social behavior and interpersonal engagement.

The evolution of social behavior

Nothing in biology makes sense except in the light of evolution. Theodosius Dobzhansky's famous dictum also holds true for explaining the evolution of love. Life on earth is fundamentally social: The ability to dynamically interact with other living organisms to support mutual homeostasis, growth, and reproduction evolved very early. Social interactions are present in primitive invertebrates and even among prokaryotes: Bacteria recognize and approach members of their own species. Bacteria also reproduce more successfully in the presence of their own kind and are able to form communities with physical and chemical characteristics that go far beyond the capabilities of the individual cell (Ingham & Ben-Jacob, 2008).

Life on earth is essentially social and many species besides humans engage in complex social interactions and have developed complex social systems that shape their behavior. [Image: www.metaphoricalplatypus.com, https://goo.gl/9n0Dli, CC BY 2.0, https://goo.gl/v4Y0Zv]

As another example, various insect species have evolved particularly complex social systems, known as eusociality. Characterized by a division of labor, eusociality appears to have evolved independently at least 11 times in insects. Research on honeybees indicates that a complex set of genes and their interactions regulate eusociality, and that these resulted from an "accelerated form of evolution" (Woodard et al., 2011). In other words, molecular mechanisms favoring high levels of sociality seem to be on an evolutionary fast track.

The evolutionary pathways that led from reptiles to mammals allowed the emergence of the unique anatomical systems and biochemical mechanisms that enable social engagement and selectively reciprocal sociality. Reptiles show minimal parental investment in offspring and form nonselective relationships between individuals. Pet owners may become emotionally attached to their turtle or snake, but this relationship is not reciprocal. In contrast, most mammals show intense parental investment in offspring and form lasting bonds with their children. Many mammalian species—including humans, wolves, and prairie voles—also develop long-lasting, reciprocal, and selective relationships between adults, with several features of what humans experience as "love." In turn, these reciprocal interactions trigger dynamic feedback mechanisms that foster growth and health.

What is love? An evolutionary and physiological perspective

Human love is more complex than simple feedback mechanisms. Love may create its own reality. The biology of love originates in the primitive parts of the brain—the emotional core of the human nervous system—which evolved long before the cerebral cortex. The brain "in love" is flooded with vague sensations, often transmitted by the **vagus nerve**, and creating much of what we experience as emotion. The modern cortex struggles to interpret love's primal messages, and weaves a narrative around incoming visceral experiences, potentially reacting to that narrative rather than to reality. It also is helpful to realize that mammalian social behavior is supported by biological components that were repurposed or co-opted over the course of mammalian evolution, eventually permitting lasting relationships between adults.

Is there a hormone of love and other relationships?

One element that repeatedly appears in the biochemistry of love is the neuropeptide **oxytocin**. In large mammals, oxytocin adopts a central role in reproduction by helping to expel the big-brained baby from the uterus, ejecting milk and sealing a selective and lasting bond between mother and offspring (Keverne, 2006). Mammalian offspring crucially depend on their mother's milk for some time after birth. Human mothers also form a strong and lasting bond with their newborns immediately after birth, in a time period that is essential for the nourishment and survival of the baby. However, women who give birth by cesarean section without going through labor, or who opt not to breastfeed, are still able to form a strong emotional bond with their children. Furthermore, fathers, grandparents, and adoptive parents also form lifelong attachments to children. Preliminary evidence suggests that the simple presence of an infant can release oxytocin in adults as well (Feldman, 2012; Kenkel et al., 2012). The baby virtually forces us to love it.

The case for a major role for oxytocin in love is strong, but until recently was based largely on extrapolation from research on parental behavior (Feldman, 2012) or social behaviors in animals (Carter, 1998; Kenkel et al., 2012). However, recent human experiments have shown that intranasal delivery of oxytocin can facilitate social behaviors, including eye contact and social cognition (Meyer-Lindenberg, Domes, Kirsch, & Heinrichs, 2011)—behaviors that are at the heart of love.

How would it feel to learn that love is nothing more than biological processes in the brain? Would it make a difference to you? Or is the fact you feel love all that matters? [Image: CC0 Public Domain, https://goo.gl/m25gce]

Of course, oxytocin is not the molecular equivalent of love. Rather, it is just one important component of a complex neurochemical system that allows the body to adapt to highly emotional situations. The systems necessary for reciprocal social interactions involve extensive neural networks through the brain and autonomic nervous system that are dynamic and constantly changing across the life span of an individual. We also now know that the properties of oxytocin are not predetermined or fixed. Oxytocin's cellular receptors are regulated by other hormones and epigenetic factors. These receptors change and adapt based on life experiences. Both oxytocin and the experience of love can change over time. In spite of limitations, new knowledge of the properties of oxytocin has proven useful in explaining several enigmatic features of love.

Stress and love

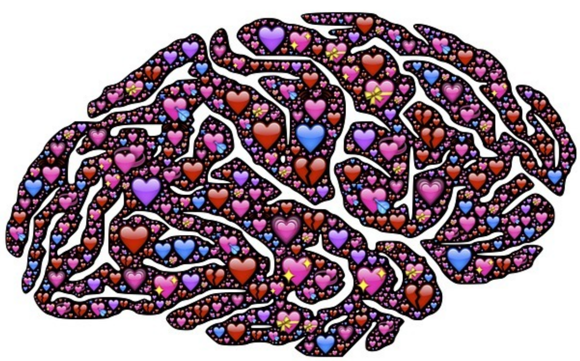
Emotional bonds can form during periods of extreme duress, especially when the survival of one individual depends on the presence and support of another. There also is evidence that oxytocin is released in response to acutely stressful experiences, perhaps serving as hormonal "insurance" against overwhelming stress. Oxytocin may help to ensure that parents and others will engage with and care for infants; develop stable, loving relationships; and seek out and receive support from others in times of need.

Animal models and the biology of social bonds

To dissect the anatomy and chemistry of love, scientists needed a biological equivalent of the

Rosetta Stone. Just as the actual stone helped linguists decipher an archaic language by comparison to a known one, animal models are helping biologists draw parallels between ancient physiology and contemporary behaviors. Studies of socially monogamous mammals that form long-lasting social bonds, such as prairie voles, have been especially helpful to an understanding the biology of human social behavior.

There is more to love than oxytocin



Just as complex as our subjective experience of love can be, so, too, is the complexity of brain processes involved with it. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Research in prairie voles showed that, as in humans, oxytocin plays a major role in social interactions and parental behavior (Carter, 1998; Carter, Boone, Pournajafi-Nazarloo, & Bales, 2009; Kenkel et al., 2012). Of course, oxytocin does not act alone. Its release and actions depend on many other neurochemicals, including endogenous opioids and dopamine (Aragona & Wang, 2009). Particularly important to social bonding are the interactions of oxytocin with a related neuropeptide known as vasopressin. The systems regulated by oxytocin and vasopressin are sometimes redundant. Both peptides are implicated in behaviors that require social engagement by either males or females, such as huddling over an infant (Kenkel et al., 2012). For example, it was necessary in voles to block both oxytocin and vasopressin receptors to induce a significant reduction in social engagement, either among adults or between adults and infants. Blocking only one of these two receptors did not eliminate social approach or contact. However, antagonists for either the oxytocin or vasopressin receptor inhibited the selective sociality, which is essential for the expression of a social bond (Bales, Kim, Lewis-Reese, & Carter, 2004; Cho, DeVries, Williams, & Carter, 1999). If we accept selective social bonds, parenting, and mate protection as proxies for love in humans, research in animals supports the hypothesis that oxytocin and vasopressin interact to allow the dynamic behavioral states and behaviors necessary for love.

Oxytocin and vasopressin have shared functions, but they are not identical in their actions. The specific behavioral roles of oxytocin and vasopressin are especially difficult to untangle because they are components of an integrated neural network with many points of intersection. Moreover, the genes that regulate the production of oxytocin and vasopressin are located on the same chromosome, possibly allowing coordinated synthesis or release of these peptides. Both peptides can bind to and have antagonist or agonist effects on each other's receptors. Furthermore, the pathways necessary for reciprocal social behavior are constantly adapting: These peptides and the systems that they regulate are always in flux. In spite of these difficulties, some of the different functions of oxytocin and vasopressin have been identified.

Functional differences between vasopressin and oxytocin

Vasopressin is associated with physical and emotional mobilization, and can help support vigilance and behaviors needed for guarding a partner or territory (Carter, 1998), as well as other forms of adaptive self-defense (Ferris, 2008). Vasopressin also may protect against physiologically "shutting down" in the face of danger. In many mammalian species, mothers exhibit agonistic behaviors in defense of their young, possibly through the interactive actions of vasopressin and oxytocin (Bosch & Neumann, 2012). Prior to mating, prairie voles are generally social, even toward strangers. However, within a day or so of mating, they begin to show high levels of aggression toward intruders (Carter, DeVries, & Getz, 1995), possibly serving to protect or guard a mate, family, or territory. This mating-induced aggression is especially obvious in males.

Oxytocin, in contrast, is associated with immobility without fear. This includes relaxed physiological states and postures that permit birth, lactation, and consensual sexual behavior. Although not essential for parenting, the increase of oxytocin associated with birth and lactation may make it easier for a woman to be less anxious around her newborn and to experience and express loving feelings for her child (Carter & Altemus, 1997). In highly social species such as prairie voles (Kenkel et al., 2013), and presumably in humans, the intricate molecular dances of oxytocin and vasopressin fine-tune the coexistence of caretaking and protective aggression.

Fatherhood also has a biological basis

The biology of fatherhood is less well-studied than motherhood is. However, male care of offspring also appears to rely on both oxytocin and vasopressin (Kenkel et al., 2012), probably acting in part through effects on the autonomic nervous system (Kenkel et al., 2013). Even sexually naïve male prairie voles show spontaneous parental behavior in the presence of an infant (Carter et al., 1995). However, the stimuli from infants or the nature of the social interactions that release oxytocin and vasopressin may differ between the sexes (Feldman, 2012).

At the heart of the benefits of love is a sense of safety

Parental care and support in a safe environment are particularly important for mental health in social mammals, including humans and prairie voles. Studies of rodents and of lactating women suggest that oxytocin has the important capacity to modulate the behavioral and autonomic distress that typically follows separation from a mother, child, or partner, reducing defensive behaviors and thereby supporting growth and health (Carter, 1998).

Love is a universal feeling of safety and connectedness with another. Research has shown that at the end of people's lives, their primary regret is not spending more time with the people they love. [Image: CC0 Public Domain, https://goo.gl/m25gce]

The absence of love in early life can be detrimental to mental and physical health

During early life in particular, trauma or neglect may produce behaviors and emotional states in humans that are socially pathological. Because the processes involved in creating social behaviors and social emotions are delicately balanced, these be may be triggered in inappropriate contexts, leading to aggression toward friends or family. Alternatively, bonds may be formed with prospective partners who fail to provide social support or protection.

Sex differences exist in the consequences of early life experiences

Males seem to be especially vulnerable to the negative effects of early experiences, possibly helping to explain the increased sensitivity of males to various developmental disorders. The implications of sex differences in the nervous system and in the response to stressful experiences for social behavior are only slowly becoming apparent (Carter et al., 2009). Both males and females produce vasopressin and oxytocin and are capable of responding to both hormones. However, in brain regions that are involved in defensive aggression, such as the extended amygdala and lateral septum, the production of vasopressin is androgen-dependent. Thus, in the face of a threat, males may be experiencing higher central levels of vasopressin.

Oxytocin and vasopressin pathways, including the peptides and their receptors, are regulated by coordinated genetic, hormonal, and epigenetic factors that influence the adaptive and behavioral functions of these peptides across the animal's life span. As a result, the endocrine and behavioral consequences of a stress or challenge may be different for males and females (DeVries, DeVries, Taymans, & Carter, 1996). For example, when unpaired prairie voles were exposed to an intense but brief stressor, such as a few minutes of swimming, or injection of the adrenal hormone corticosterone, the males (but not females) quickly formed new pair bonds. These and other experiments suggest that males and females have different coping strategies, and possibly may experience both stressful experiences, and even love, in ways that are gender-specific.

In the context of nature and evolution, sex differences in the nervous system are important. However, sex differences in brain and behavior also may help to explain gender differences in the vulnerability to mental and physical disorders (Taylor, et al., 2000). Better understanding these differences will provide clues to the physiology of human mental health in both sexes.

Loving relationships in early life can have epigenetic consequences

Although we are all born with a finite set of genes, experiences in childhood will cause some genes to express themselves (e.g., encourage certain personality traits), while other genes will remain dormant. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Love is "epigenetic." That is, positive experiences in early life can act upon and alter the expression of specific genes. These changes in gene expression may have behavioral consequences through simple biochemical changes, such as adding a methyl group to a particular site within the genome (Zhang & Meaney, 2010). It is possible that these changes in the genome may even be passed to the next generation.

Social behaviors, emotional attachment to others, and long-lasting reciprocal relationships also are both plastic and adaptive, and so is the biology upon which they are based. For example, infants of traumatized or highly stressed parents might be chronically exposed to vasopressin, either through their own increased production of the peptide, or through higher levels of vasopressin in maternal milk. Such increased exposure could sensitize the infant to defensive behaviors or create a lifelong tendency to overreact to threat. Based on research in rats, it seems that in response to adverse early experiences of chronic isolation, the genes for vasopressin receptors can become upregulated (Zhang et al., 2012), leading to an increased sensitivity to acute stressors or anxiety that may persist throughout life.

Epigenetic programming triggered by early life experiences is adaptive in allowing neuroendocrine systems to project and plan for future behavioral demands. But epigenetic changes that are long-lasting also can create atypical social or emotional behaviors (Zhang & Meaney, 2010) that may be especially likely to surface in later life, and in the face of social or emotional challenges.

Exposure to exogenous hormones in early life also may be epigenetic. For example, prairie voles treated postnatally with vasopressin (especially males) were later more aggressive, whereas those exposed to a vasopressin antagonist showed less aggression in adulthood. Conversely, in voles the exposure of infants to slightly increased levels of oxytocin during development increased the tendency to show a pair bond. However, these studies also showed that a single exposure to a higher level of oxytocin in early life could disrupt the later capacity to pair bond (Carter et al., 2009).

There is little doubt that either early social experiences or the effects of developmental exposure to these neuropeptides holds the potential to have long-lasting effects on behavior. Both parental care and exposure to oxytocin in early life can permanently modify hormonal systems, altering the capacity to form relationships and influence the expression of love across the life span. Our preliminary findings in voles further suggest that early life experiences affect the methylation of the oxytocin receptor gene and its expression (Connelly, Kenkel, Erickson, & Carter, 2011). Thus, we can plausibly argue that love is epigenetic.

The absence of social behavior or isolation also has consequences for the oxytocin system

Given the power of positive social experiences, it is not surprising that a lack of social relationships also may lead to alterations in behavior as well as changes in oxytocin and vasopressin pathways. We have found that social isolation reduced the expression of the gene for the oxytocin receptor, and at the same time increased the expression of genes for the vasopressin peptide. In female prairie voles, isolation also was accompanied by an increase in blood levels of oxytocin, possibly as a coping mechanism. However, over time, isolated prairie voles of both sexes showed increases in measures of depression, anxiety, and physiological arousal, and these changes were observed even when endogenous oxytocin was elevated. Thus, even the hormonal insurance provided by endogenous oxytocin in face of the chronic stress of isolation was not sufficient to dampen the consequences of living alone. Predictably, when isolated voles were given additional exogenous oxytocin, this treatment did restore many of these functions to normal (Grippo, Trahanas, Zimmerman, Porges, & Carter, 2009).

In modern societies, humans can survive, at least after childhood, with little or no human contact. Communication technology, social media, electronic parenting, and many other recent technological advances may reduce social behaviors, placing both children and adults at risk for social isolation and disorders of the autonomic nervous system, including deficits in their capacity for social engagement and love (Porges, 2011).

Social engagement actually helps us to cope with stress. The same hormones and areas of the brain that increase the capacity of the body to survive stress also enable us to better adapt to an ever-changing social and physical environment. Individuals with strong emotional support and relationships are more resilient in the face of stressors than those who feel isolated or lonely. Lesions in various bodily tissues, including the brain, heal more quickly in animals that are living socially versus in isolation (Karelina & DeVries, 2011). The protective effects of positive sociality seem to rely on the same cocktail of hormones that carries a biological message of “love” throughout the body.

Can love—or perhaps oxytocin—be a medicine?

Although research has only begun to examine the physiological effects of these peptides beyond social behavior, there is a wealth of new evidence showing that oxytocin can influence physiological responses to stress and injury. As only one example, the molecules associated with love have restorative properties, including the ability to literally heal a “broken heart.” Oxytocin receptors are expressed in the heart, and precursors for oxytocin appear to be critical for the development of the fetal heart (Danalache, Gutkowska, Slusarz, Berezowska, & Jankowski, 2010). Oxytocin exerts protective and restorative effects in part through its capacity to convert undifferentiated stem cells into cardiomyocytes. Oxytocin can facilitate adult neurogenesis and tissue repair, especially after a stressful experience. We now know that oxytocin has direct anti-inflammatory and antioxidant properties in *in vitro* models of atherosclerosis (Szeto et al., 2008). The heart seems to rely on oxytocin as part of a normal process of protection and self-healing.

Researchers are interested in the medical/therapeutic potential of oxytocin. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Thus, oxytocin exposure early in life not only regulates our ability to love and form social bonds, it also affects our health and well-being. Oxytocin modulates the hypothalamic–pituitary adrenal (HPA) axis, especially in response to disruptions in homeostasis (Carter, 1998), and coordinates demands on the immune system and energy balance. Long-term, secure relationships provide emotional support and down-regulate reactivity of the HPA axis, whereas intense stressors, including birth, trigger activation of the HPA axis and sympathetic nervous system. The ability of oxytocin to regulate these systems probably explains the exceptional capacity of most women to cope with the challenges of childbirth and childrearing.

Dozens of ongoing clinical trials are currently attempting to examine the therapeutic potential of oxytocin in disorders ranging from autism to heart disease. Of course, as in hormonal studies in voles, the effects are likely to depend on the history of the individual and the context, and to be dose-dependent. As this research is emerging, a variety of individual differences and apparent discrepancies in the effects of exogenous oxytocin are being reported. Most of these studies do not include any information on the endogenous hormones, or on the oxytocin or vasopressin receptors, which are likely to affect the outcome of such treatments.

Conclusion

Research in this field is new and there is much left to understand. However, it is already clear that both love and oxytocin are powerful. Of course, with power comes responsibility. Although research into mechanisms through which love—or hormones such as oxytocin—may protect us against stress and disease is in its infancy, this knowledge will ultimately increase our understanding of the way that our emotions impact upon health and disease. The same molecules that allow us to give and receive love also link our need for others with health and well-being.

Acknowledgments

C. Sue Carter and Stephen W. Porges are both Professors of Psychiatry at the University of North Carolina, Chapel Hill, and also are Research Professors of Psychology at Northeastern University, Boston.

Discussions of "love and forgiveness" with members of the Fetzer Institute's Advisory Committee on Natural Sciences led to this essay and are gratefully acknowledged here. We are especially appreciative of thoughtful editorial input from Dr. James Harris. Studies from the authors' laboratories were sponsored by the National Institutes of Health. We also express our gratitude for this support and to our colleagues, whose input and hard work informed the ideas expressed in this article. A version of this paper was previously published in *EMBO Reports* in the series on "Sex and Society"; this paper is reproduced with the permission of the publishers of that journal.

Outside Resources

Book: C. S. Carter, L. Ahnert et al. (Eds.), (2006). Attachment and bonding: A new synthesis. Cambridge, MA: MIT Press.

Book: Porges, S.W. (2011). The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication and self-regulation. New York, NY: Norton.

Web: Database of publicly and privately supported clinical studies of human participants conducted around the world.
http://www.clinicaltrials.gov

Web: PubMed comprises over 22 million citations for biomedical literature from MEDLINE, life science journals, and online books. PubMed citations and abstracts include the fields of biomedicine and health, covering portions of the life sciences, behavioral sciences, chemical sciences, and bioengineering. PubMed also provides access to additional relevant web sites and links to the other NCBI molecular biology resources.
http://www.ncbi.nlm.nih.gov/pubmed

Web: Website of author Stephen Porges
http://www.stephenporges.com/

Discussion Questions

1. If love is so important in human behavior, why is it so hard to describe and understand?
2. Discuss the role of evolution in understanding what humans call “love” or other forms of prosociality.
3. What are the common biological and neuroendocrine elements that appear in maternal love and adult-adult relationships?
4. Oxytocin and vasopressin are biochemically similar. What are some of the differences between the actions of oxytocin and vasopressin?
5. How may the properties of oxytocin and vasopressin help us understand the biological bases of love?
6. What are common features of the biochemistry of “love” and “safety,” and why are these important to human health?

Vocabulary

Epigenetics
Heritable changes in gene activity that are not caused by changes in the DNA sequence. http://en.wikipedia.org/wiki/Epigenetics

Oxytocin
A nine amino acid mammalian neuropeptide. Oxytocin is synthesized primarily in the brain, but also in other tissues such as uterus, heart and thymus, with local effects. Oxytocin is best known as a hormone of female reproduction due to its capacity to cause uterine contractions and eject milk. Oxytocin has effects on brain tissue, but also acts throughout the body in some cases as an antioxidant or anti-inflammatory.

Vagus nerve
The 10th cranial nerve. The mammalian vagus has an older unmyelinated branch which originates in the dorsal motor complex and a more recently evolved, myelinated branch, with origins in the ventral vagal complex including the nucleus ambiguous. The vagus is the primary source of autonomic-parasympathetic regulation for various internal organs, including the heart, lungs and other parts of the viscera. The vagus nerve is primarily sensory (afferent), transmitting abundant visceral input to the central nervous system.

Vasopressin
A nine amino acid mammalian neuropeptide. Vasopressin is synthesized primarily in the brain, but also may be made in other tissues. Vasopressin is best known for its effects on the cardiovascular system (increasing blood pressure) and also the kidneys (causing water retention). Vasopressin has effects on brain tissue, but also acts throughout the body.

References

Aragona, B. J, & Wang, Z. (2009). Dopamine regulation of social choice in a monogamous rodent species. *Frontiers in Behavioral Neuroscience, 3*, 15.

Bales, K. L., Kim, A. J., Lewis-Reese, A. D., & Carter, C. S. (2004). Both oxytocin and vasopressin may influence alloparental care in male prairie voles. *Hormones and Behavior, 44*, 454–361.

Bosch, O. J., & Neumann, I. D. (2012). Both oxytocin and vasopressin are mediators of maternal care and aggression in rodents: from central release to sites of action. *Hormones and Behavior, 61*, 293–303.

Carter, C. S. (1998). Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology, 23*, 779–818.

Carter, C. S., & Altemus, M. (1997). Integrative functions of lactational hormones in social behavior and stress management. *Annals of the New York Academy of Sciences, Integrative Neurobiology of Affiliation 807*, 164–174.

Carter, C. S., Boone, E. M., Pournajafi-Nazarloo, H., & Bales, K. L. (2009). The consequences of early experiences and exposure to oxytocin and vasopressin are sexually-dimorphic. *Developmental Neuroscience, 31*, 332–341.

Carter, C. S., DeVries, A. C., & Getz, L. L. (1995). Physiological substrates of mammalian monogamy: The prairie vole model. *Neuroscience and Biobehavioral Reviews, 19*, 303–314.

Cho, M. M., DeVries, A. C., Williams, J. R., Carter, C. S. (1999). The effects of oxytocin and vasopressin on partner preferences in male and female prairie voles (Microtus ochrogaster). *Behavioral Neuroscience, 113*, 1071–1080.

Connelly, J., Kenkel, W., Erickson, E., & Carter, C. S. (2011). Are birth and oxytocin epigenetic events. *Society for Neuroscience Abstracts, 388*.10.

Danalache, B. A., Gutkowska, J., Slusarz, M. J., Berezowska, I., & Jankowski, M. (2010). Oxytocin-Gly-Lys-Arg: A novel cardiomyogenic peptide. *PloS One, 5*(10), e13643.

DeVries, A. C., DeVries, M. B., Taymans, S. E., & Carter, C. S. (1996). Stress has sexually dimorphic effects on pair bonding in prairie voles. *Proceedings of the National Academy of Sciences, USA, 93*, 11980–11984.

Feldman, R. (2012). Oxytocin and social affiliation in humans. *Hormones and Behavior, 61*, 380–391.

Ferris, C. F. (2008). Functional magnetic resonance imaging and the neurobiology of vasopressin and oxytocin. *Progress in Brain Research, 170*, 305–320.

Grippo, A. J., Trahanas, D. M., Zimmerman, R. R., Porges, S. W., & Carter, C. S. (2009). Oxytocin protects against negative behavioral and autonomic consequences of long-term social isolation. *Psychoneuroendocrinology, 34*, 1542–1553.

Ingham, C. J., & Ben-Jacob, E. (2008). Swarming and complex pattern formation in Paenicbachillus vortex studied by imaging and tracking cells. *BMC Microbiol*, 8, 36.

Karelina, K., & DeVries, A. C. (2011). Modeling social influences on human health. *Psychosomatic Medicine, 73*, 67–74.

Kenkel, W.M., Paredes, J., Lewis, G.F., Yee, J.R., Pournajafi-Nazarloo, H., Grippo, A.J., Porges, S. W., & Carter, C.S. (2013). Autonomic substrates of the response to pups in male prairie voles. *PlosOne* Aug 5;8(8):e69965. doi: 10.1371/journal.pone.0069965.

Kenkel, W.M., Paredes, J., Yee, J. R., Pournajafi-Nazarloo, H., Bales, K. L., & Carter, C. S. (2012). Exposure to an infant releases oxytocin and facilitates pair-bonding in male prairie voles. *Journal of Neuroendocrinology, 24*, 874–886.

Keverne, E. B. (2006). Neurobiological and molecular approaches to attachment and bonding. In C. S. Carter, L. Ahnert et al. (Eds.), *Attachment and bonding: A new synthesis*. Cambridge, MA: MIT Press. Pp. 101-117.

Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature: Reviews in Neuroscience 12*, 524–538.

Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication and self-regulation*. New York, NY: Norton.

Szeto, A., Nation, D. A., Mendez, A. J., Dominguez-Bendala, J., Brooks, L. G., Schneiderman, N., & McCabe, P. M. (2008). Oxytocin attenuates NADPH-dependent superoxide activity and IL-6 secretion in macrophages and vascular cells. *American Journal of Physiology, Endocrinology and Metabolism 295*, E1495–1501.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychol Rev, 107*, 411–429.

Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron S.A., . . . Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc Natl Acad Sci USA, 108*, 7472–7477.

Zhang, L., Hernandez, V. S., Liu, B., Medina, M. P., Nava-Kopp, A. T., Irles, C., & Morales, M. (2012). Hypothalamic vasopressin system regulation by maternal separation: Its impact on anxiety in rats. *Neuroscience, 215*, 135–148.

Zhang, T. Y., & Meaney, M. J. (2010). Epigenetics and the environmental regulation of the genome and its function. *Annual Review of Psychology, 61*, 439–466.

10
Epigenetics in Psychology

Ian Weaver

Early life experiences exert a profound and long-lasting influence on physical and mental health throughout life. The efforts to identify the primary causes of this have significantly benefited from studies of the epigenome—a dynamic layer of information associated with DNA that differs between individuals and can be altered through various experiences and environments. The epigenome has been heralded as a key "missing piece" of the etiological puzzle for understanding how development of psychological disorders may be influenced by the surrounding environment, in concordance with the genome. Understanding the mechanisms involved in the initiation, maintenance, and heritability of epigenetic states is thus an important aspect of research in current biology, particularly in the study of learning and memory, emotion, and social behavior in humans. Moreover, epigenetics in psychology provides a framework for understanding how the expression of genes is influenced by experiences and the environment to produce individual differences in behavior, cognition, personality, and mental health. In this module, we survey recent developments revealing epigenetic aspects of mental health and review some of the challenges of epigenetic approaches in psychology to help explain how nurture shapes nature.

Learning Objectives

- Explain what the term epigenetics means and the molecular machinery involved.
- Name and discuss important neural and developmental pathways that are regulated by epigenetic factors, and provide examples of epigenetic effects on personality traits and cognitive behavior.
- Understand how misregulation of epigenetic mechanisms can lead to disease states, and be able to discuss examples.
- Recognize how epigenetic machinery can be targets for therapeutic agents, and discuss

examples.

Introduction

DNA stands for Deoxyribonucleic Acid, and although each person's DNA is unique to that individual, it is 99.9% similar to every other human on the planet. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Early childhood is not only a period of physical growth; it is also a time of mental development related to changes in the anatomy, physiology, and chemistry of the nervous system that influence mental health throughout life. Cognitive abilities associated with learning and memory, reasoning, problem solving, and developing relationships continue to emerge during childhood. Brain development is more rapid during this critical or sensitive period than at any other, with more than 700 neural connections created each second. Herein, complex gene–environment interactions (or genotype–environment interactions, G×E) serve to increase the number of possible contacts between neurons, as they hone their adult synaptic properties and excitability. Many weak connections form to different neuronal targets; subsequently, they undergo remodeling in which most connections vanish and a few stable connections remain. These structural changes (or plasticity) may be crucial for the development of mature neural networks that support emotional, cognitive, and social behavior. The generation of different morphology, physiology, and behavioral outcomes from a single genome in response to changes in the environment forms the basis for "phenotypic plasticity," which is fundamental to the way organisms cope with environmental variation, navigate the present world, and solve future problems.

The challenge for psychology has been to integrate findings from genetics and environmental (social, biological, chemical) factors, including the quality of infant–mother attachments, into the study of personality and our understanding of the emergence of mental illness. These studies have demonstrated that common DNA sequence variation and rare mutations account for only a small fraction (1%–2%) of the total risk for inheritance of personality traits and mental disorders (Dick, Riley, & Kendler, 2010; Gershon, Alliey-Rodriguez, & Liu, 2011). Additionally, studies that have attempted to examine the mechanisms and conditions under which DNA sequence variation influences brain development and function have been confounded by complex cause-and-effect relationships (Petronis, 2010). The large unaccounted heritability of personality traits and mental health suggests that additional molecular and cellular mechanisms are involved.

Epigenetics has the potential to provide answers to these important questions and refers to the transmission of **phenotype** in terms of gene expression in the absence of changes in DNA sequence—hence the name epi- (Greek: επί- over, above) genetics (Waddington, 1942; Wolffe & Matzke, 1999). The advent of high-throughput techniques such as sequencing-based approaches to study the distributions of regulators of gene expression throughout the genome led to the collective description of the "epigenome." In contrast to the genome sequence, which is static and the same in almost all cells, the **epigenome** is highly dynamic, differing among cell types, tissues, and brain regions (Gregg et al., 2010). Recent studies have provided insights into epigenetic regulation of developmental pathways in response to a range of external environmental factors (Dolinoy, Weidman, & Jirtle, 2007). These environmental factors during early childhood and adolescence can cause changes in expression of genes conferring risk of mental health and chronic physical conditions. Thus, the examination of genetic–epigenetic–environment interactions from a developmental perspective may determine the nature of gene misregulation in psychological disorders.

This module will provide an overview of the main components of the epigenome and review themes in recent epigenetic research that have relevance for psychology, to form the biological basis for the interplay between environmental signals and the genome in the regulation of individual differences in physiology, emotion, cognition, and behavior.

Molecular control of gene expression: the dynamic epigenome

Almost all the cells in our body are genetically identical, yet our body generates many different cell types, organized into different tissues and organs, and expresses different proteins. Within each type of mammalian cell, about 2 meters of genomic DNA is divided into nuclear chromosomes. Yet the nucleus of a human cell, which contains the chromosomes, is only about 2 μm in diameter. To achieve this 1,000,000-fold compaction, DNA is wrapped around a group of 8 proteins called histones. This combination of DNA and histone proteins forms a special structure called a "nucleosome," the basic unit of chromatin, which represents a structural solution for maintaining and accessing the tightly compacted genome. These factors

Identical twins are the perfect example of epigenetics. Although they share exactly the same DNA, their unique experiences in life will cause some genes (and not others) to express themselves. This is why, over time, identical twins come to look and behave differently. [Image: M., https://goo.gl/VU5iJv, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

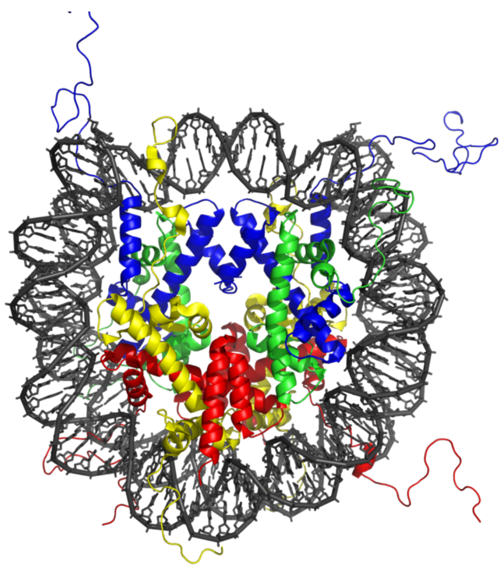
alter the likelihood that a gene will be expressed or silenced. Cellular functions such as gene expression, DNA replication, and the generation of specific cell types are therefore influenced by distinct patterns of chromatin structure, involving covalent modification of both histones (Kadonaga, 1998) and DNA (Razin, 1998).

Importantly, epigenetic variation also emerges across the lifespan. For example, although identical twins share a common genotype and are genetically identical and epigenetically similar when they are young, as they age they become more dissimilar in their epigenetic patterns and often display behavioral, personality, or even physical differences, and have different risk levels for serious illness. Thus, understanding the structure of the nucleosome is key to understanding the precise and stable control of gene expression and regulation, providing a molecular interface between genes and environmentally induced changes in cellular activity.

The primary epigenetic mark: DNA modification

DNA methylation is the best-understood epigenetic modification influencing gene expression. DNA is composed of four types of naturally occurring nitrogenous bases: adenine (A), thymine (T), guanine (G), and cytosine (C). In mammalian genomes, DNA methylation occurs primarily at cytosine residues in the context of cytosines that are followed by guanines (CpG dinucleotides), to form 5-methylcytosine in a cell-specific pattern (Goll & Bestor, 2005; Law & Jacobsen, 2010; Suzuki & Bird, 2008). The enzymes that perform DNA methylation are called DNA methyltransferases (DNMTs), which catalyze the transfer of a methyl group to the cytosine (Adams, McKay, Craig, & Burdon, 1979). These enzymes are all expressed in the central nervous system and are dynamically regulated during development (Feng, Chang, Li, & Fan, 2005; Goto et al., 1994). The effect of DNA methylation on gene function varies depending on the period of development during which the methylation occurs and location of the methylated cytosine. Methylation of DNA in gene regulatory regions (promoter and enhancer regions) usually results in gene silencing and reduced gene expression (Ooi, O'Donnell, & Bestor, 2009; Suzuki & Bird, 2008; Sutter and Doerfler, 1980; Vardimon et al., 1982). This is a powerful regulatory mechanism that ensures that genes are expressed only when needed. Thus DNA methylation may broadly impact human brain development, and age-related misregulation of DNA methylation is associated with the molecular pathogenesis of neurodevelopmental disorders.

Histone modification and the histone code



Life experiences, like a stressful event in childhood, can cause the modification of histone proteins (pictured) to help adapt to one's environment. For example, in response to a stressful event, histone modification of one's DNA might occur to encourage a more cautions personality—in order to avoid future, stressful encounters. [Image: Zephyris, https://goo.gl/gGrSQd, CC BY-SA 3.0, https://goo.gl/kB1Ogc]

The modification of histone proteins comprises an important epigenetic mark related to gene expression. One of the most thoroughly studied modifications is histone acetylation, which is associated with gene activation and increased gene expression (Wade, Pruss, & Wolffe, 1997). Acetylation on histone tails is mediated by the opposing enzymatic activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs) (Kuo & Allis, 1998). For example, acetylation of histone in gene regulatory regions by HAT enzymes is generally associated with DNA demethylation, gene activation, and increased gene expression (Hong, Schroth, Matthews, Yau, & Bradbury, 1993; Sealy & Chalkley, 1978). On the other hand, removal of the acetyl group (deacetylation) by HDAC enzymes is generally associated with DNA methylation, gene silencing, and decreased gene expression (Davie & Chadee, 1998). The relationship between patterns of histone modifications and gene activity provides evidence for the existence of a "histone code" for determining cell-specific gene expression programs (Jenuwein & Allis, 2001). Interestingly, recent research using animal models has demonstrated that histone modifications and DNA methylation of certain genes mediates the long-term behavioral effects of the level of care experienced during infancy.

Early childhood experience

The development of an individual is an active process of adaptation that occurs within a social and economic context. For example, the closeness or degree of positive attachment of the parent (typically mother)–infant bond and parental investment (including nutrient supply provided by the parent) that define early childhood experience also program the development of individual differences in stress responses in the brain, which then affect memory, attention, and emotion. In terms of evolution, this process provides the offspring with the ability to physiologically adjust gene expression profiles contributing to the organization and function of neural circuits and molecular pathways that support (1) biological defensive systems for survival (e.g., stress resilience), (2) reproductive success to promote establishment and persistence in the present environment, and (3) adequate parenting in the next generation (Bradshaw, 1965).

Parental investment and programming of stress responses in the offspring

The most comprehensive study to date of variations in parental investment and epigenetic inheritance in mammals is that of the maternally transmitted responses to stress in rats. In rat pups, maternal nurturing (licking and grooming) during the first week of life is associated with long-term programming of individual differences in stress responsiveness, emotionality, cognitive performance, and reproductive behavior (Caldji et al., 1998; Francis, Diorio, Liu, & Meaney, 1999; Liu et al., 1997; Myers, Brunelli, Shair, Squire, & Hofer, 1989; Stern, 1997). In adulthood, the offspring of mothers that exhibit increased levels of pup licking and grooming over the first week of life show increased expression of the glucocorticoid receptor in the hippocampus (a brain structure associated with stress responsivity as well as learning and memory) and a lower hormonal response to stress compared with adult animals reared by low licking and grooming mothers (Francis et al., 1999; Liu et al., 1997). Moreover, rat pups that received low levels of maternal licking and grooming during the first week of life showed decreased histone acetylation and increased DNA methylation of a neuron-specific promoter of the glucocorticoid receptor gene (Weaver et al., 2004). The expression of this gene is then reduced, the number of glucocorticoid receptors in the brain is decreased, and the animals show a higher hormonal response to stress throughout their life. The effects of maternal care on stress hormone responses and behaviour in the offspring can be eliminated in adulthood by pharmacological treatment (HDAC inhibitor trichostatin A, TSA) or dietary amino acid supplementation (methyl donor L-methionine), treatments that influence histone acetylation, DNA methylation, and expression of the glucocorticoid receptor gene (Weaver et al., 2004; Weaver et al., 2005). This series of experiments shows that histone acetylation and DNA methylation of the glucocorticoid receptor gene promoter is a necessary link in the process leading to the long-term physiological and behavioral sequelae of poor maternal care. This points to a possible molecular target for treatments that may reverse or ameliorate the traces of childhood maltreatment.

Several studies have attempted to determine to what extent the findings from model animals are transferable to humans. Examination of post-mortem brain tissue from healthy human subjects found that the human equivalent of the glucocorticoid receptor gene promoter (NR3C1 exon 1F promoter) is also unique to the individual (Turner, Pelascini, Macedo, & Muller, 2008). A similar study examining newborns showed that methylation of the glucocorticoid receptor gene promoter maybe an early epigenetic marker of maternal mood and risk of increased hormonal responses to stress in infants 3 months of age (Oberlander et al., 2008). Although further studies are required to examine the functional consequence of this DNA methylation, these findings are consistent with our studies in the neonate and adult offspring of low licking and grooming mothers that show increased DNA methylation of the promoter of the glucocorticoid receptor gene, decreased glucocorticoid receptor gene expression, and increased hormonal responses to stress (Weaver et al., 2004). Examination of brain tissue from suicide victims found that the human glucocorticoid receptor gene promoter is also more methylated in the brains of individuals who had experienced maltreatment during childhood (McGowan et al., 2009). These finding suggests that DNA methylation mediates the effects of early environment in both rodents and humans and points to the possibility of new therapeutic approaches stemming from translational epigenetic research. Indeed, similar processes at comparable epigenetic labile regions could explain why the adult offspring of high and low licking/grooming mothers exhibit widespread differences in hippocampal gene expression and cognitive function (Weaver, Meaney, & Szyf, 2006).

Parental care during one's childhood has important and consequential effects on the development of an individual, effects that persist even into adulthood. [Image: The White Ribbon Alliance, https://goo.gl/KgY6N5, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

However, this type of research is limited by the inaccessibility of human brain samples. The translational potential of this finding would be greatly enhanced if the relevant epigenetic modification can be measured in an accessible tissue. Examination of blood samples from adult patients with bipolar disorder, who also retrospectively reported on their experiences of childhood abuse and neglect, found that the degree of DNA methylation of the human glucocorticoid receptor gene promoter was strongly positively related to the reported experience of childhood maltreatment decades earlier. For a relationship between a molecular measure and reported historical exposure, the effects size is extraordinarily large. This opens a range of new possibilities: given the large effect size and consistency of this association, measurement of the GR promoter methylation may effectively become a blood test measuring the physiological traces left on the genome by early experiences. Although this blood test cannot replace current methods of diagnosis, this unique and addition information adds to our knowledge of how disease may arise and be manifested throughout life. Near-future research will examine whether this measure adds value over and above simple reporting of early adversities when it comes to predicting important outcomes, such as response to treatment or suicide.

Child nutrition and the epigenome

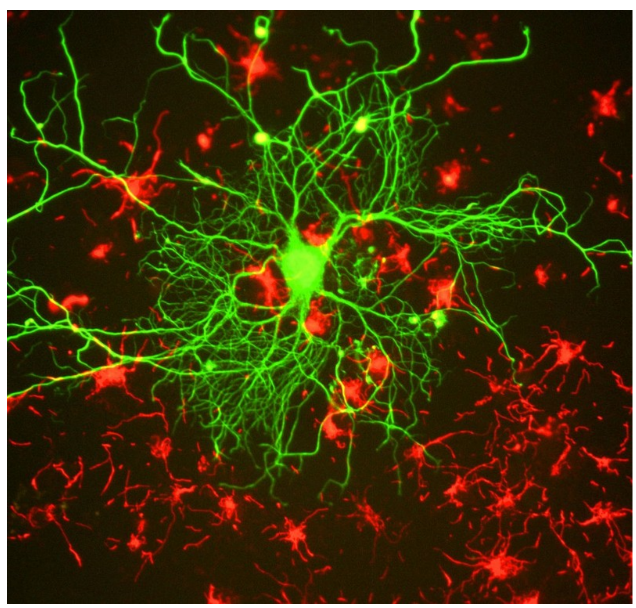
Whether or not your parents knew the science behind it, telling you to eat your veggies as a kid really does make you healthier and stronger—at least your DNA, that is. [Image: U.S. Department of Agriculture, https://goo.gl/tpyYzA, CC BY 2.0, https://goo.gl/BRvSA7]

The old adage "you are what you eat" might be true on more than just a physical level: The food you choose (and even what your parents and grandparents chose) is reflected in your own personal development and risk for disease in adult life (Wells, 2003). Nutrients can reverse or change DNA methylation and histone modifications, thereby modifying the expression of critical genes associated with physiologic and pathologic processes, including embryonic development, aging, and carcinogenesis. It appears that nutrients can influence the epigenome either by directly inhibiting enzymes that catalyze DNA methylation or histone modifications, or by altering the availability of substrates necessary for those enzymatic reactions. For example, rat mothers fed a diet low in methyl group donors during pregnancy produce offspring with reduced DNMT-1 expression, decreased DNA methylation, and increased histone acetylation at promoter regions of specific genes, including the glucocorticoid receptor, and increased gene expression in the liver of juvenile offspring (Lillycrop, Phillips, Jackson, Hanson, & Burdge, 2005) and adult offspring (Lillycrop et al., 2007). These data suggest that early life nutrition has the potential to influence epigenetic programming in the brain not only during early development but also in adult life, thereby modulating health throughout life. In this regard, nutritional epigenetics has been viewed as an attractive tool to prevent pediatric developmental diseases and cancer, as well as to delay aging-associated processes.

The best evidence relating to the impact of adverse environmental conditions development and health comes from studies of the children of women who were pregnant during two civilian famines of World War II: the Siege of Leningrad (1941–44) (Bateson, 2001) and the Dutch Hunger Winter (1944–1945) (Stanner et al., 1997). In the Netherlands famine, women who were previously well nourished were subjected to low caloric intake and associated environmental stressors. Women who endured the famine in the late stages of pregnancy gave birth to smaller babies (Lumey & Stein, 1997) and these children had an increased risk of insulin resistance later in life (Painter, Roseboom, & Bleker, 2005). In addition, offspring who were starved prenatally later experienced impaired glucose tolerance in adulthood, even when food was more abundant (Stanner et al., 1997). Famine exposure at various stages of gestation was associated with a wide range of risks such as increased obesity, higher rates of coronary heart disease, and lower birth weight (Lumey & Stein, 1997). Interestingly, when examined 60 years later, people exposed to famine prenatally showed reduced DNA methylation compared with their unexposed same-sex siblings (Heijmans et al., 2008).

Epigenetic regulation of learning and memory

Memories are recollections of actual events stored within our brains. But how is our brain able to form and store these memories? Epigenetic mechanisms influence genomic activities in the brain to produce long-term changes in synaptic signaling, organization, and morphology, which in turn support learning and memory (Day & Sweatt, 2011).



Neural plasticity is the change of neural pathways and synapses which allows for our ability to learn new things and remember them. [Image: Gerry Shaw, https://goo.gl/JBqlY7, CC BY-SA 3.0, https://goo.gl/eLCn2O]

Neuronal activity in the hippocampus of mice is associated with changes in DNA methylation (Guo et al., 2011), and disruption to genes encoding the DNA methylation machinery cause learning and memory impairments (Feng et al., 2010). DNA methylation has also been implicated in the maintenance of long-term memories, as pharmacological inhibition of DNA methylation and impaired memory (Day & Sweatt, 2011; Miller et al., 2010). These findings indicate the importance of DNA methylation in mediating synaptic plasticity and cognitive functions, both of which are disturbed in psychological illness.

Changes in histone modifications can also influence long-term memory formation by altering chromatin accessibility and the expression of genes relevant to learning and memory. Memory formation and the associated enhancements in synaptic transmission are accompanied by increases in histone acetylation (Guan et al., 2002) and alterations in histone methylation (Schaefer et al., 2009), which promote gene expression. Conversely, a neuronal increase in histone deacetylase activity, which promotes gene silencing, results in reduced synaptic plasticity and impairs memory (Guan et al., 2009). Pharmacological inhibition of histone deacetylases augments memory formation (Guan et al., 2009; Levenson et al., 2004), further suggesting that histone (de)acetylation regulates this process.

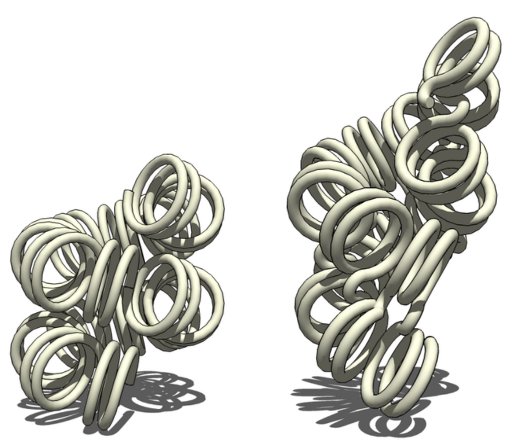
In humans genetic defects in genes encoding the DNA methylation and chromatin machinery exhibit profound effects on cognitive function and mental health (Jiang, Bressler, & Beaudet, 2004). The two best-characterized examples are Rett syndrome (Amir et al., 1999) and Rubinstein-Taybi syndrome (RTS) (Alarcon et al., 2004), which are profound intellectual disability disorders. Both MECP2 and CBP are highly expressed in neurons and are involved in regulating neural gene expression (Chen et al., 2003; Martinowich et al., 2003).

Rett syndrome patients have a mutation in their DNA sequence in a gene called MECP2. MECP2 plays many important roles within the cell: One of these roles is to read the DNA sequence, checking for DNA methylation, and to bind to areas that contain methylation, thereby preventing the wrong proteins from being present. Other roles for MECP2 include promoting the presence of particular, necessary, proteins, ensuring that DNA is packaged properly within the cell and assisting with the production of proteins. MECP2 function also influences gene expression that supports dendritic and synaptic development and hippocampus-dependent memory (Li, Zhong, Chau, Williams, & Chang, 2011; Skene et al., 2010). Mice with altered MECP2 expression exhibit genome-wide increases in histone acetylation, neuron cell death, increased anxiety, cognitive deficits, and social withdrawal (Shahbazian et al., 2002). These findings support a model in which DNA methylation and MECP2 constitute a cell-specific epigenetic mechanism for regulation of histone modification and gene expression, which may be disrupted in Rett syndrome.

RTS patients have a mutation in their DNA sequence in a gene called CBP. One of these roles of CBP is to bind to specific histones and promote histone acetylation, thereby promoting gene expression. Consistent with this function, RTS patients exhibit a genome-wide decrease in histone acetylation and cognitive dysfunction in adulthood (Kalkhoven et al., 2003). The learning and memory deficits are attributed to disrupted neural plasticity (Korzus, Rosenfeld, & Mayford, 2004). Similar to RTS in humans, mice with a mutation of CBP perform poorly in cognitive tasks and show decreased genome-wide histone acetylation (for review, see Josselyn, 2005). In the mouse brain CBP was found to act as an epigenetic switch to promote the birth of new neurons in the brain. Interestingly, this epigenetic mechanism is disrupted in the fetal brains of mice with a mutation of CBP, which, as pups, exhibit early behavioral deficits following removal and separation from their mother (Wang et al., 2010). These findings provide a novel mechanism whereby environmental cues, acting through histone modifying enzymes, can regulate epigenetic status and thereby directly promote neurogenesis, which regulates neurobehavioral development.

Together, these studies demonstrate that misregulation of epigenetic modifications and their regulatory enzymes is capable of orchestrating prominent deficits in neuronal plasticity and cognitive function. Knowledge from these studies may provide greater insight into other mental disorders such as depression and suicidal behaviors.

Epigenetic mechanisms in psychological disorders



Pictured above is a chromatin, the spiral-looking macromolecule involved in depression. [Image: Zephyris, https://goo.gl/6DBQ1g, CC BY-SA 3.0, https://goo.gl/eLCn2O]

Epigenome-wide studies have identified several dozen sites with DNA methylation alterations in genes involved in brain development and neurotransmitter pathways, which had previously been associated with mental illness (Mill et al., 2008). These disorders are complex and typically start at a young age and cause lifelong disability. Often, limited benefits from treatment make these diseases some of the most burdensome disorders for individuals, families, and society. It has become evident that the efforts to identify the primary causes of complex psychiatric disorders may significantly benefit from studies linking environmental effects with changes observed within the individual cells.

Epigenetic events that alter chromatin structure to regulate programs of gene expression have been associated with depression-related behavior and action of antidepressant medications, with increasing evidence for similar mechanisms occurring in post-mortem brains of depressed individuals. In mice, social avoidance resulted in decreased expression of hippocampal genes important in mediating depressive responses (Tsankova et al., 2006). Similarly, chronic social defeat stress was found to decrease expression of genes implicated in normal emotion processing (Lutter et al., 2008). Consistent with these findings, levels of histone markers of increased gene expression were down regulated in human post-mortem brain samples from individuals with a history of clinical depression (Covington et al., 2009).

Administration of antidepressants increased histone markers of increased gene expression and reversed the gene repression induced by defeat stress (Lee, Wynder, Schmidt, McCafferty, & Shiekhattar, 2006; Tsankova et al., 2006; Wilkinson et al., 2009). These results provide support for the use of HDAC inhibitors against depression. Accordingly, several HDAC inhibitors have been found to exert antidepressant effects by each modifying distinct cellular targets (Cassel et al., 2006; Schroeder, Lin, Crusio, & Akbarian, 2007).

There is also increasing evidence that aberrant gene expression resulting from altered epigenetic regulation is associated with the pathophysiology of suicide (McGowan et al., 2008; Poulter et al., 2008). Thus, it is tempting to speculate that there is an epigenetically determined reduced capacity for gene expression, which is required for learning and memory, in the brains of suicide victims.

Although there is some evidence that a dysfunctional upbringing can increase one's likelihood for schizophrenia (an epigenetically inherited disease), some people who have both the predisposition and the stressful environment never develop the mental illness. [Image: Steve White, CC0 Public Domain, https://goo.gl/m25gce]

Epigenetic strategy to understanding gene-environment interactions

While the cellular and molecular mechanisms that influence on physical and mental health have long been a central focus of neuroscience, only in recent years has attention turned to the epigenetic mechanisms behind the dynamic changes in gene expression responsible for normal cognitive function and increased risk for mental illness. The links between early environment and epigenetic modifications suggest a mechanism underlying gene-environment interactions. Early environmental adversity alone is not a sufficient cause of mental illness, because many individuals with a history of severe childhood maltreatment or trauma remain healthy. It is increasingly becoming evident that inherited differences in the segments of specific genes may moderate the effects of adversity and determine who is sensitive and who is resilient through a gene-environment interplay. Genes such as the glucocorticoid receptor appear to moderate the effects of childhood adversity on mental illness. Remarkably, epigenetic DNA modifications have been identified that may underlie the long-lasting effects of environment on biological functions. This new epigenetic research is pointing to a new strategy to understanding gene-environment interactions.

The next decade of research will show if this potential can be exploited in the development of new therapeutic options that may alter the traces that early environment leaves on the genome. However, as discussed in this module, the epigenome is not static and can be molded by developmental signals, environmental perturbations, and disease states, which present an experimental challenge in the search for epigenetic risk factors in psychological disorders (Rakyan, Down, Balding, & Beck, 2011). The sample size and epigenomic assay required is dependent on the number of tissues affected, as well as the type and distribution of epigenetic modifications. The combination of genetic association maps studies with epigenome-wide developmental studies may help identify novel molecular mechanisms to explain features of inheritance of personality traits and transform our understanding of the biological basis of psychology. Importantly, these epigenetic studies may lead to identification of novel therapeutic targets and enable the development of improved strategies for early diagnosis, prevention, and better treatment of psychological and behavioral disorders.

Outside Resources

Reference: The "Encyclopedia of DNA Elements" (ENCODE) project
http://encodeproject.org/ENCODE/

Reference: THREADS - A new way to explore the ENCODE Project
http://www.nature.com/encode/#/threads

Web: Explore, view, and download genome-wide maps of DNA and histone modifications from the NCBI Epigenomics Portal
http://www.ncbi.nlm.nih.gov/epigenomics

Web: NOVA ScienceNOW - Introduction to Epigenetics
http://www.pbs.org/wgbh/nova/genes

Web: The University of Utah's Genetic Science Learning Center
http://learn.genetics.utah.edu/content/epigenetics/

Discussion Questions

1. Describe the physical state of the genome when genes are active and inactive.
2. Often, the physical characteristics of genetically identical twins become increasingly different as they age, even at the molecular level. Explain why this is so (use the terms "environment" and "epigenome").
3. Name 3–4 environmental factors that influence the epigenome and describe their effects.
4. The rat nurturing example shows us how parental behavior can shape the behavior of offspring on a biochemical level. Discuss how this relates to humans and include the personal and social implications.
5. Explain how the food we eat affects gene expression.
6. Can the diets of parents affect their offspring's epigenome?
7. Why is converging evidence the best kind of evidence in the study of brain function?
8. If you were interested in whether a particular brain area was involved in a specific behavior, what neuroscience methods could you use?
9. If you were interested in the precise time in which a particular brain process occurred,

which neuroscience methods could you use?

Vocabulary

DNA methylation
Covalent modifications of mammalian DNA occurring via the methylation of cytosine, typically in the context of the CpG dinucleotide.

DNA methyltransferases (DNMTs)
Enzymes that establish and maintain DNA methylation using methyl-group donor compounds or cofactors. The main mammalian DNMTs are DNMT1, which maintains methylation state across DNA replication, and DNMT3a and DNMT3b, which perform de novo methylation.

Epigenetics
The study of heritable changes in gene expression or cellular phenotype caused by mechanisms other than changes in the underlying DNA sequence. Epigenetic marks include covalent DNA modifications and posttranslational histone modifications.

Epigenome
The genome-wide distribution of epigenetic marks.

Gene
A specific deoxyribonucleic acid (DNA) sequence that codes for a specific polypeptide or protein or an observable inherited trait.

Genome-wide association study (GWAS)
A study that maps DNA polymorphisms in affected individuals and controls matched for age, sex, and ethnic background with the aim of identifying causal genetic variants.

Genotype
The DNA content of a cell's nucleus, whether a trait is externally observable or not.

Histone acetyltransferases (HATs) and histone deacetylases (HDACs)
HATs are enzymes that transfer acetyl groups to specific positions on histone tails, promoting an "open" chromatin state and transcriptional activation. HDACs remove these acetyl groups, resulting in a "closed" chromatin state and transcriptional repression.

Histone modifications
Posttranslational modifications of the N-terminal "tails" of histone proteins that serve as a major mode of epigenetic regulation. These modifications include acetylation,

phosphorylation, methylation, sumoylation, ubiquitination, and ADP-ribosylation.

Identical twins

Two individual organisms that originated from the same zygote and therefore are genetically identical or very similar. The epigenetic profiling of identical twins discordant for disease is a unique experimental design as it eliminates the DNA sequence-, age-, and sex-differences from consideration.

Phenotype

The pattern of expression of the genotype or the magnitude or extent to which it is observably expressed—an observable characteristic or trait of an organism, such as its morphology, development, biochemical or physiological properties, or behavior.

References

Adams, R. L., McKay, E. L., Craig, L. M., & Burdon, R. H. (1979). Mouse DNA methylase: methylation of native DNA. *Biochimica et Biophysica Acta, 561*(2), 345–357.

Alarcon, J. M., Malleret, G., Touzani, K., Vronskaya, S., Ishii, S., Kandel, E. R., & Barco, A. (2004). Chromatin acetylation, memory, and LTP are impaired in CBP+/- mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. *Neuron, 42*(6), 947–959. doi: 10.1016/j.neuron.2004.05.021, S0896627304003022 [pii]

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics, 23*(2), 185–188.

Bateson, P. (2001). Fetal experience and good adult design. *International Journal of Epidemiology, 30*(5), 928–934.

Bradshaw, A.D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics, 13*, 115–155.

Caldji, C., Tannenbaum, B., Sharma, S., Francis, D., Plotsky, P. M., & Meaney, M. J. (1998). Maternal care during infancy regulates the development of neural systems mediating the expression of fearfulness in the rat. *Proceedings of the National Academy of Sciences U S A, 95*(9), 5335–5340.

Cassel, S., Carouge, D., Gensburger, C., Anglard, P., Burgun, C., Dietrich, J. B., . . . Zwiller, J. (2006). Fluoxetine and cocaine induce the epigenetic factors MeCP2 and MBD1 in adult rat brain. *Molecular Pharmacology, 70*(2), 487–492. doi: 10.1124/mol.106.022301

Chen, W. G., Chang, Q., Lin, Y., Meissner, A., West, A. E., Griffith, E. C., . . . Greenberg, M. E. (2003). Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science, 302*(5646), 885–889.

Covington, H. E., 3rd, Maze, I., LaPlant, Q. C., Vialou, V. F., Ohnishi, Y. N., Berton, O., . . . Nestler, E. J. (2009). Antidepressant actions of histone deacetylase inhibitors. *Journal of Neuroscience, 29*(37), 11451–11460. doi: 10.1523/JNEUROSCI.1758-09.2009

Davie, J. R., & Chadee, D. N. (1998). Regulation and regulatory parameters of histone modifications. *Journal of Cellular Biochemistry Suppl, 30–31 *, 203–213.

Day, J. J., & Sweatt, J. D. (2011). Epigenetic mechanisms in cognition. *Neuron, 70*(5), 813–829. doi: 10.1016/j.neuron.2011.05.019

Dick, D. M., Riley, B., & Kendler, K. S. (2010). Nature and nurture in neuropsychiatric genetics: where do we stand? *Dialogues in Clinical Neuroscience, 12*(1), 7–23.

Dolinoy, D. C., Weidman, J. R., & Jirtle, R. L. (2007). Epigenetic gene regulation: linking early developmental environment to adult disease. *Reproductive Toxicology, 23*(3), 297–307. doi: S0890-6238(06)00197-3 [pii], 10.1016/j.reprotox.2006.08.012

Feng, J., Chang, H., Li, E., & Fan, G. (2005). Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system. *Journal of Neuroscience Research, 79*(6), 734–746. doi: 10.1002/jnr.20404

Feng, J., Zhou, Y., Campbell, S. L., Le, T., Li, E., Sweatt, J. D., . . . Fan, G. (2010). Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nature Neuroscience, 13*(4), 423–430. doi: 10.1038/nn.2514

Francis, D., Diorio, J., Liu, D., & Meaney, M. J. (1999). Nongenomic transmission across generations of maternal behavior and stress responses in the rat. *Science, 286*(5442), 1155–1158.

Gershon, E. S., Alliey-Rodriguez, N., & Liu, C. (2011). After GWAS: searching for genetic risk for schizophrenia and bipolar disorder. *American Journal of Psychiatry, 168*(3), 253–256. doi: 10.1176/appi.ajp.2010.10091340

Goll, M. G., & Bestor, T. H. (2005). Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry, 74*, 481–514. doi: 10.1146/annurev.biochem.74.010904.153721

Goto, K., Numata, M., Komura, J. I., Ono, T., Bestor, T. H., & Kondo, H. (1994). Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice. *Differentiation, 56*(1–2), 39–44.

Gregg, C., Zhang, J., Weissbourd, B., Luo, S., Schroth, G. P., Haig, D., & Dulac, C. (2010). High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science, 329* (5992), 643–648. doi: 10.1126/science.1190830

Guan, J. S., Haggarty, S. J., Giacometti, E., Dannenberg, J. H., Joseph, N., Gao, J., . . . Tsai, L. H. (2009). HDAC2 negatively regulates memory formation and synaptic plasticity. *Nature, 459* (7243), 55-60. doi: 10.1038/nature07925

Guan, Z., Giustetto, M., Lomvardas, S., Kim, J. H., Miniaci, M. C., Schwartz, J. H., . . . Kandel, E. R. (2002). Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. *Cell, 111*(4), 483–493.

Guo, J. U., Ma, D. K., Mo, H., Ball, M. P., Jang, M. H., Bonaguidi, M. A., . . . Song, H. (2011). Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nature Neuroscience, 14*(10), 1345–1351. doi: 10.1038/nn.2900

Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., . . . Lumey, L. H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences U S A, 105*(44), 17046–17049. doi:

0806560105 [pii]10.1073/pnas.0806560105

Hong, L., Schroth, G. P., Matthews, H. R., Yau, P., & Bradbury, E. M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *Journal of Biological Chemistry, 268*(1), 305–314.

Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. *Science, 293*(5532), 1074–1080. doi: 10.1126/Science.1063127293/5532/1074 [pii]

Jiang, Y. H., Bressler, J., & Beaudet, A. L. (2004). Epigenetics and human disease. *Annual Review of Genomics and Human Genetics, 5*, 479–510. doi: 10.1146/annurev.genom.5.061903.180014

Josselyn, S. A. (2005). What's right with my mouse model? New insights into the molecular and cellular basis of cognition from mouse models of Rubinstein-Taybi Syndrome. *Learning & Memory, 12*(2), 80–83. doi: 12/2/80 [pii]10.1101/lm.93505

Kadonaga, J. T. (1998). Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell, 92*(3), 307–313.

Kalkhoven, E., Roelfsema, J. H., Teunissen, H., den Boer, A., Ariyurek, Y., Zantema, A., . . . Peters, D. J. (2003). Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome. *Human Molecular Genetics, 12*(4), 441–450.

Korzus, E., Rosenfeld, M. G., & Mayford, M. (2004). CBP histone acetyltransferase activity is a critical component of memory consolidation. *Neuron, 42*(6), 961–972. doi: 10.1016/j.neuron.2004.06.002S0896627304003526 [pii]

Kuo, M. H., & Allis, C. D. (1998). *Roles of histone acetyltransferases and deacetylases in gene regulation*. Bioessays, 20(8), 615–626. doi: 10.1002/(SICI)1521–1878(199808)20:8<615::AID-BIES4>3.0.CO;2-H [pii] 10.1002/(SICI)1521-1878(199808)20:8<615::AID-BIES4>3.0.CO;2-H

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics, 11*(3), 204–220. doi: nrg2719 [pii] 10.1038/nrg2719

Lee, M. G., Wynder, C., Schmidt, D. M., McCafferty, D. G., & Shiekhattar, R. (2006). Histone H3 lysine 4 demethylation is a target of nonselective antidepressive medications. *Chemistry & Biology, 13*(6), 563–567. doi: 10.1016/j.chembiol.2006.05.004

Levenson, J. M., O'Riordan, K. J., Brown, K. D., Trinh, M. A., Molfese, D. L., & Sweatt, J. D. (2004). Regulation of histone acetylation during memory formation in the hippocampus. *Journal of Biological Chemistry, 279*(39), 40545–40559.

Li, H., Zhong, X., Chau, K. F., Williams, E. C., & Chang, Q. (2011). Loss of activity-induced phosphorylation of MeCP2 enhances synaptogenesis, LTP and spatial memory. *Nature Neuroscience, 14*(8), 1001–1008. doi: 10.1038/nn.2866

Lillycrop, K. A., Phillips, E. S., Jackson, A. A., Hanson, M. A., & Burdge, G. C. (2005). Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. *Journal of Nutrition, 135*(6), 1382–1386. doi: 135/6/1382 [pii]

Lillycrop, K. A., Slater-Jefferies, J. L., Hanson, M. A., Godfrey, K. M., Jackson, A. A., & Burdge, G. C. (2007). Induction of altered epigenetic regulation of the hepatic glucocorticoid receptor in the offspring of rats fed a protein-restricted diet during pregnancy suggests that reduced DNA methyltransferase-1 expression is involved in impaired DNA methylation and changes in histone modifications. *British Journal of Nutrition, 97*(6), 1064–1073. doi: S000711450769196X [pii]10.1017/S000711450769196X

Liu, D., Diorio, J., Tannenbaum, B., Caldji, C., Francis, D., Freedman, A., . . . Meaney, M. J. (1997). Maternal care, hippocampal glucocorticoid receptors, and hypothalamic- pituitary-adrenal responses to stress [see comments]. *Science, 277*(5332), 1659–1662.

Lumey, L. H., & Stein, A. D. (1997). Offspring birth weights after maternal intrauterine undernutrition: a comparison within sibships. *American Journal of Epidemiology, 146*(10), 810–819.

Lutter, M., Krishnan, V., Russo, S. J., Jung, S., McClung, C. A., & Nestler, E. J. (2008). Orexin signaling mediates the antidepressant-like effect of calorie restriction. *Journal of Neuroscience, 28*(12), 3071–3075. doi: 10.1523/JNEUROSCI.5584-07.2008

Martinowich, K., Hattori, D., Wu, H., Fouse, S., He, F., Hu, Y., . . . Sun, Y. E. (2003). DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science, 302* (5646), 890–893.

McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonte, B., Szyf, M., . . . Meaney, M. J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience, 12*(3), 342–348. doi: nn.2270 [pii]10.1038/nn.2270

McGowan, P. O., Sasaki, A., Huang, T. C., Unterberger, A., Suderman, M., Ernst, C., . . . Szyf, M. (2008). Promoter-wide hypermethylation of the ribosomal RNA gene promoter in the suicide brain. *PLoS ONE, 3*(5), e2085. doi: 10.1371/journal.pone.0002085

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., . . . Petronis, A. (2008). Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *American Journal of Human Genetics, 82*(3), 696–711. doi: 10.1016/j.ajhg.2008.01.008

Miller, C. A., Gavin, C. F., White, J. A., Parrish, R. R., Honasoge, A., Yancey, C. R., . . . Sweatt, J. D. (2010). Cortical DNA methylation maintains remote memory. *Nature Neuroscience, 13*(6), 664–666. doi: 10.1038/nn.2560

Myers, M. M., Brunelli, S. A., Shair, H. N., Squire, J. M., & Hofer, M. A. (1989). Relationships between maternal behavior of SHR and WKY dams and adult blood pressures of cross-fostered F1 pups. *Developmental Psychobiology, 22*(1), 55–67.

Oberlander, T. F., Weinberg, J., Papsdorf, M., Grunau, R., Misri, S., & Devlin, A. M. (2008). Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics, 3*(2), 97–106. doi: 6034 [pii]

Ooi, S. K., O'Donnell, A. H., & Bestor, T. H. (2009). Mammalian cytosine methylation at a glance. *Journal of Cell Science, 122*(Pt 16), 2787–2791. doi: 122/16/2787 [pii]10.1242/jcs.015123

Painter, R. C., Roseboom, T. J., & Bleker, O. P. (2005). Prenatal exposure to the Dutch famine and disease in later life: an overview. *Reproductive Toxicology, 20*(3), 345–352. doi: S0890-6238(05)00088-2 [pii]10.1016/j. Reproductive Toxicology.2005.04.005

Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature, 465*(7299), 721–727. doi: 10.1038/nature09230

Poulter, M. O., Du, L., Weaver, I. C., Palkovits, M., Faludi, G., Merali, Z., . . . Anisman, H. (2008). GABAA receptor promoter hypermethylation in suicide brain: implications for the involvement of epigenetic processes. *Biological Psychiatry, 64*(8), 645–652. doi: 10.1016/j. biopsych.2008.05.028

Rakyan, V. K., Down, T. A., Balding, D. J., & Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics, 12*(8), 529–541. doi: 10.1038/nrg3000

Razin, A. (1998). CpG methylation, chromatin structure and gene silencing-a three-way connection. *European Molecular Biology Organization, 17*(17), 4905–4908.

Schaefer, A., Sampath, S. C., Intrator, A., Min, A., Gertler, T. S., Surmeier, D. J., . . . Greengard, P. (2009). Control of cognition and adaptive behavior by the GLP/G9a epigenetic suppressor complex. *Neuron, 64*(5), 678–691. doi: 10.1016/j.neuron.2009.11.019

Schroeder, F. A., Lin, C. L., Crusio, W. E., & Akbarian, S. (2007). Antidepressant-like effects of the histone deacetylase inhibitor, sodium butyrate, in the mouse. *Biological Psychiatry, 62* (1), 55-64. doi: 10.1016/j.biopsych.2006.06.036

Sealy, L., & Chalkley, R. (1978). DNA associated with hyperacetylated histone is preferentially digested by DNase I. *Nucleic Acids Research, 5*(6), 1863–1876.

Shahbazian, M., Young, J., Yuva-Paylor, L., Spencer, C., Antalffy, B., Noebels, J., . . . Zoghbi, H. (2002). Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron, 35*(2), 243–254.

Skene, P. J., Illingworth, R. S., Webb, S., Kerr, A. R., James, K. D., Turner, D. J., . . . Bird, A. P. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Molecular Cell, 37*(4), 457–468. doi: 10.1016/j.molcel.2010.01.030

Stanner, S. A., Bulmer, K., Andres, C., Lantseva, O. E., Borodina, V., Poteen, V. V., & Yudkin, J. S. (1997). Does malnutrition in utero determine diabetes and coronary heart disease in adulthood? Results from the Leningrad siege study, a cross sectional study. *British Medical Journal, 315*(7119), 1342–1348.

Stern, J. M. (1997). Offspring-induced nurturance: animal-human parallels. *Developmental Psychobiololgy, 31*(1), 19–37.

Sutter, D., Doerfler, W., 1980. Methylation of integrated adenovirus type 12 DNA sequences in transformed cells is inversely correlated with viral gene expression. *Proceedings of the National Academy of Sciences U S A. 77*, 253–256.

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics, 9*(6), 465–476. doi: nrg2341 [pii]10.1038/nrg2341

Tsankova, N. M., Berton, O., Renthal, W., Kumar, A., Neve, R. L., & Nestler, E. J. (2006). Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nature Neuroscience. 9*(4): 519–525. doi:10.1038/nn1659

Turner, J. D., Pelascini, L. P., Macedo, J. A., & Muller, C. P. (2008). Highly individual methylation patterns of alternative glucocorticoid receptor promoters suggest individualized epigenetic regulatory mechanisms. *Nucleic Acids Research, 36*(22), 7207–7218. doi: gkn897 [pii] 10.1093/nar/gkn897

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M., Doerfler, W., 1982. Expression of a cloned adenovirus gene is inhibited by in vitro methylation. *Proceedings of the National Academy of Sciences U S A. 79*, 1073–1077.

Waddington, C. H. (1942). Epigenotype. *Endeavour*(1), 18–21.

Wade, P. A., Pruss, D., & Wolffe, A. P. (1997). Histone acetylation: chromatin in action. *Trends in Biochemical Sciences, 22*(4), 128–132. doi: S0968000497010165 [pii]

Wang, J., Weaver, I. C., Gauthier-Fisher, A., Wang, H., He, L., Yeomans, J., . . . Miller, F. D. (2010). CBP histone acetyltransferase activity regulates embryonic neural differentiation in the normal and Rubinstein-Taybi syndrome brain. *Developmental Cell, 18*(1), 114–125. doi: 10.1016/j.devcel.2009.10.023

Weaver, I. C., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., . . . Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience, 7*(8), 847–854. doi: 10.1038/nn1276

Weaver, I. C., Champagne, F. A., Brown, S. E., Dymov, S., Sharma, S., Meaney, M. J., & Szyf, M. (2005). Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: altering epigenetic marking later in life. *Journal of Neuroscience, 25*(47), 11045–11054. doi: 10.1523/JNEUROSCI.3652-05.2005

Weaver, I. C., Meaney, M. J., & Szyf, M. (2006). Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. *Proceedings of the National Academy of Sciences U S A, 103*(9), 3480–3485. doi: 10.1073/pnas.0507526103

Wells, J. C. (2003). The thrifty phenotype hypothesis: thrifty offspring or thrifty mother? *Journal of Theoretical Biology, 221*(1), 143–161.

Wilkinson, M. B., Xiao, G., Kumar, A., LaPlant, Q., Renthal, W., Sikder, D., . . . Nestler, E. J. (2009). Imipramine treatment and resiliency exhibit similar chromatin regulation in the mouse nucleus accumbens in depression models. *Journal of Neuroscience, 29*(24), 7820–7832. doi: 10.1523/JNEUROSCI.0932-09.2009

Wolffe, A. P., & Matzke, M. A. (1999). Epigenetics: regulation through repression. *Science, 286* (5439), 481–486.

11
The Nature-Nurture Question

Eric Turkheimer

People have a deep intuition about what has been called the “nature–nurture question.” Some aspects of our behavior feel as though they originate in our genetic makeup, while others feel like the result of our upbringing or our own hard work. The scientific field of behavior genetics attempts to study these differences empirically, either by examining similarities among family members with different degrees of genetic relatedness, or, more recently, by studying differences in the DNA of people with different behavioral traits. The scientific methods that have been developed are ingenious, but often inconclusive. Many of the difficulties encountered in the empirical science of behavior genetics turn out to be conceptual, and our intuitions about nature and nurture get more complicated the harder we think about them. In the end, it is an oversimplification to ask how “genetic” some particular behavior is. Genes and environments always combine to produce behavior, and the real science is in the discovery of how they combine for a given behavior.

Learning Objectives

- Understand what the nature–nurture debate is and why the problem fascinates us.
- Understand why nature–nurture questions are difficult to study empirically.
- Know the major research designs that can be used to study nature–nurture questions.
- Appreciate the complexities of nature–nurture and why questions that seem simple turn out not to have simple answers.

Introduction

There are three related problems at the intersection of philosophy and science that are fundamental to our understanding of our relationship to the natural world: the mind–body problem, the free will problem, and the nature–nurture problem. These great questions have a lot in common. Everyone, even those without much knowledge of science or philosophy, has opinions about the answers to these questions that come simply from observing the world we live in. Our feelings about our relationship with the physical and biological world often seem incomplete. We are in control of our actions in some ways, but at the mercy of our bodies in others; it feels obvious that our consciousness is some kind of creation of our physical brains, at the same time we sense that our awareness must go beyond just the physical. This incomplete knowledge of our relationship with nature leaves us fascinated and a little obsessed, like a cat that climbs into a paper bag and then out again, over and over, mystified every time by a relationship between inner and outer that it can see but can't quite understand.

It may seem obvious that we are born with certain characteristics while others are acquired, and yet of the three great questions about humans' relationship with the natural world, only nature–nurture gets referred to as a "debate." In the history of psychology, no other question has caused so much controversy and offense: We are so concerned with nature–nurture because our very sense of moral character seems to depend on it. While we may admire the athletic skills of a great basketball player, we think of his height as simply a gift, a payoff in the "genetic lottery." For the same reason, no one blames a short person for his height or someone's congenital disability on poor decisions: To state the obvious, it's "not their fault." But we do praise the concert violinist (and perhaps her parents and teachers as well) for her dedication, just as we condemn cheaters, slackers, and bullies for their bad behavior.

The problem is, most human characteristics aren't usually as clear-cut as height or instrument-mastery, affirming our nature–nurture expectations strongly one way or the other. In fact, even the great violinist might have some inborn qualities—perfect pitch, or long, nimble fingers—that support and reward her hard work. And the basketball player might have eaten a diet while growing up that promoted his genetic tendency for being tall. When we think about our own qualities, they seem under our control in some respects, yet beyond our control in others. And often the traits that don't seem to have an obvious cause are the ones that concern us the most and are far more personally significant. What about how much we drink or worry? What about our honesty, or religiosity, or sexual orientation? They all come from that uncertain zone, neither fixed by nature nor totally under our own control.

One major problem with answering nature-nurture questions about people is, how do you set up an experiment? In nonhuman animals, there are relatively straightforward experiments for tackling nature–nurture questions. Say, for example, you are interested in aggressiveness

Researchers have learned a great deal about the nature-nurture dynamic by working with animals. But of course many of the techniques used to study animals cannot be applied to people. Separating these two influences in human subjects is a greater research challenge. [Image: Sebastián Dario, https://goo.gl/OPiIWd, CC BY-NC 2.0, https://goo.gl/FIlc2e]

in dogs. You want to test for the more important determinant of aggression: being born to aggressive dogs or being raised by them. You could mate two aggressive dogs—angry Chihuahuas—together, and mate two nonaggressive dogs—happy beagles—together, then switch half the puppies from each litter between the different sets of parents to raise. You would then have puppies born to aggressive parents (the Chihuahuas) but being raised by nonaggressive parents (the Beagles), and vice versa, in litters that mirror each other in puppy distribution. The big questions are: Would the Chihuahua parents raise aggressive beagle puppies? Would the beagle parents raise *non*aggressive Chihuahua puppies? Would the puppies' *nature* win out, regardless of who raised them? Or… would the result be a combination of nature *and* nurture? Much of the most significant nature–nurture research has been done in this way (Scott & Fuller, 1998), and animal breeders have been doing it successfully for thousands of years. In fact, it is fairly easy to breed animals for behavioral traits.

With people, however, we can't assign babies to parents at random, or select parents with certain behavioral characteristics to mate, merely in the interest of science (though history does include horrific examples of such practices, in misguided attempts at "eugenics," the shaping of human characteristics through intentional breeding). In typical human families, children's biological parents raise them, so it is very difficult to know whether children act like their parents due to genetic (nature) or environmental (nurture) reasons. Nevertheless, despite our restrictions on setting up human-based experiments, we do see real-world examples of nature-nurture at work in the human sphere—though they only provide partial answers to our many questions.

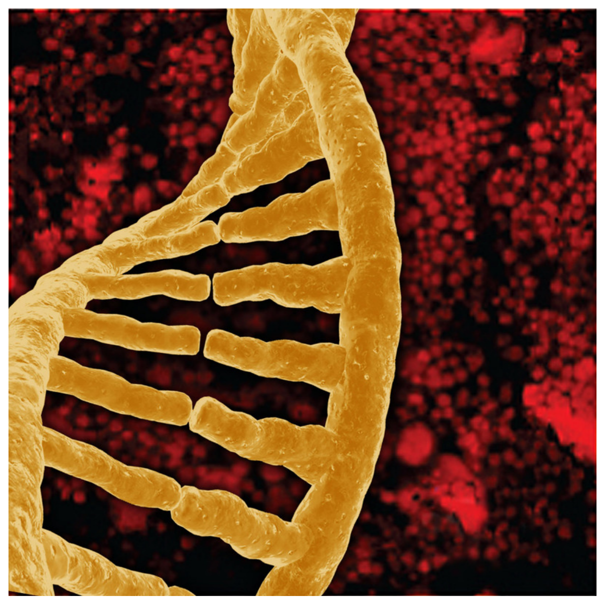
The science of how genes and environments work together to influence behavior is called behavioral genetics. The easiest opportunity we have to observe this is the adoption study. When children are put up for adoption, the parents who give birth to them are no longer the parents who raise them. This setup isn't quite the same as the experiments with dogs (children aren't assigned to random adoptive parents in order to suit the particular interests of a scientist) but adoption still tells us some interesting things, or at least confirms some basic expectations. For instance, if the biological child of tall parents were adopted into a family of short people, do you suppose the child's growth would be affected? What about the biological child of a Spanish-speaking family adopted at birth into an English-speaking family? What language would you expect the child to speak? And what might these outcomes tell you about the difference between height and language in terms of nature-nurture?

Another option for observing nature-nurture in humans involves **twin studies**. There are two types of twins: monozygotic (MZ) and dizygotic (DZ). Monozygotic twins, also called "identical" twins, result from a single zygote (fertilized egg) and have the same DNA. They are essentially clones. Dizygotic twins, also known as "fraternal" twins, develop from two zygotes and share 50% of their DNA. Fraternal twins are ordinary siblings who happen to have been born at the same time. To analyze nature–nurture using twins, we compare the similarity of MZ and DZ pairs. Sticking with the features of height and spoken language, let's take a look at how nature and nurture apply: Identical twins, unsurprisingly, are almost perfectly similar for height. The heights of fraternal twins, however, are like any other sibling pairs: more similar to each other than to people from other families, but hardly identical. This contrast between twin types gives us a clue about the role genetics plays in determining height. Now consider spoken language. If one identical twin speaks Spanish at home, the co-twin with whom she is raised almost certainly does too. But the same would be true for a pair of fraternal twins raised together. In terms of spoken language, fraternal twins are just as similar as identical twins, so it appears that the genetic match of identical twins doesn't make much difference.

Studies focused on twins have led to important insights about the biological origins of many personality characteristics.

Twin and adoption studies are two instances of a much broader class of methods for observing nature-nurture called **quantitative genetics**, the scientific discipline in which similarities among individuals are analyzed based on how biologically related they are. We can do these studies with siblings and half-siblings, cousins, twins who have been separated at birth and raised separately (Bouchard, Lykken, McGue, & Segal, 1990; such twins are very rare and play a smaller role than is commonly believed in the science of nature–nurture), or with entire extended families (see Plomin, DeFries, Knopik, & Neiderhiser, 2012, for a complete introduction to research methods relevant to nature–nurture).

For better or for worse, contentions about nature–nurture have intensified because quantitative genetics produces a number called a **heritability coefficient**, varying from 0 to 1, that is meant to provide a single measure of genetics' influence of a trait. In a general way, a heritability coefficient measures how strongly differences among individuals are related to differences among their genes. But beware: Heritability coefficients, although simple to compute, are deceptively difficult to interpret. Nevertheless, numbers that provide simple answers to complicated questions tend to have a strong influence on the human imagination, and a great deal of time has been spent discussing whether the heritability of intelligence or personality or depression is equal to one number or another.



Quantitative genetics uses statistical methods to study the effects that both heredity and environment have on test subjects. These methods have provided us with the heritability coefficient which measures how strongly differences among individuals for a trait are related to differences among their genes. [Image: EMSL, https://goo.gl/IRfn9g, CC BY-NC-SA 2.0, https://goo.gl/fbv27n]

One reason nature–nurture continues to fascinate us so much is that we live in an era of great scientific discovery in genetics, comparable to the times of Copernicus, Galileo, and Newton, with regard to astronomy and physics. Every day, it seems, new discoveries are made, new possibilities proposed. When Francis Galton first started thinking about nature–nurture in the late-19th century he was very influenced by his cousin, Charles Darwin, but genetics *per se* was unknown. Mendel's famous work with peas, conducted at about the same time, went undiscovered for 20 years; quantitative genetics was developed in the 1920s; DNA was discovered by Watson and Crick in the 1950s; the human genome was completely sequenced at the turn of the 21st century; and we are now on the verge of being able to obtain the specific DNA sequence of anyone at a relatively low cost. No one knows what this new genetic knowledge will mean for the study of nature–nurture, but as we will see in the next section, answers to nature–nurture questions have turned out to be far more difficult and mysterious than anyone imagined.

What Have We Learned About Nature–Nurture?

It would be satisfying to be able to say that nature–nurture studies have given us conclusive and complete evidence about where traits come from, with some traits clearly resulting from genetics and others almost entirely from environmental factors, such as childrearing practices and personal will; but that is not the case. Instead, *everything* has turned out to have some footing in genetics. The more genetically-related people are, the more similar they are—for *everything*: height, weight, intelligence, personality, mental illness, etc. Sure, it seems like common sense that some traits have a genetic bias. For example, adopted children resemble their biological parents even if they have never met them, and identical twins are more similar to each other than are fraternal twins. And while certain psychological traits, such as personality or mental illness (e.g., schizophrenia), seem reasonably influenced by genetics, it turns out that the same is true for political attitudes, how much television people watch (Plomin, Corley, DeFries, & Fulker, 1990), and whether or not they get divorced (McGue & Lykken, 1992).

Research over the last half century has revealed how central genetics are to behavior. The more genetically related people are the more similar they are not just physically but also in terms of personality and behavior. [Image: Paul Altobelli, https://goo.gl/SWLwm2, CC BY 2.0, https://goo.gl/9uSnqN]

It may seem surprising, but genetic influence on behavior is a relatively recent discovery. In the middle of the 20th century, psychology was dominated by the doctrine of behaviorism, which held that behavior could only be explained in terms of environmental factors. Psychiatry concentrated on psychoanalysis, which probed for roots of behavior in individuals' early life-histories. The truth is, neither behaviorism nor psychoanalysis is incompatible with genetic influences on behavior, and neither Freud nor Skinner was naive about the importance of organic processes in behavior. Nevertheless, in their day it was widely thought that children's personalities were shaped entirely by imitating their parents' behavior, and that schizophrenia was caused by certain kinds of "pathological mothering." Whatever the outcome of our broader discussion of nature–nurture, the basic fact that the best predictors of an adopted child's personality or mental health are found in the biological parents he or she has never met, rather than in the adoptive parents who raised him or her, presents a significant challenge to purely environmental explanations of personality or psychopathology. The message is clear: You can't leave genes out of the equation. But keep in mind, no behavioral traits are completely inherited, so you can't leave the environment out altogether, either.

Trying to untangle the various ways nature-nurture influences human behavior can be messy, and often common-sense notions can get in the way of good science. One very significant contribution of behavioral genetics that has changed psychology for good can be very helpful to keep in mind: When your subjects are biologically-related, no matter how clearly a situation may seem to point to environmental influence, it is never safe to interpret a behavior as wholly the result of nurture without further evidence. For example, when presented with data showing that children whose mothers read to them often are likely to have better reading scores in third grade, it is tempting to conclude that reading to your kids out loud is important to success in school; this may well be true, but the study as described is inconclusive, because there are genetic *as well as* environmental pathways between the parenting practices of mothers and the abilities of their children. This is a case where "correlation does not imply causation," as they say. To establish that reading aloud causes success, a scientist can either study the problem in adoptive families (in which the genetic pathway is absent) or by finding a way to randomly assign children to oral reading conditions.

The outcomes of nature–nurture studies have fallen short of our expectations (of establishing clear-cut bases for traits) in many ways. The most disappointing outcome has been the inability to organize traits from *more*- to *less*-genetic. As noted earlier, everything has turned out to be at least *somewhat* heritable (passed down), yet nothing has turned out to be *absolutely* heritable, and there hasn't been much consistency as to which traits are *more* heritable and which are *less* heritable once other considerations (such as how accurately the trait can be measured) are taken into account (Turkheimer, 2000). The problem is conceptual: The heritability coefficient, and, in fact, the whole quantitative structure that underlies it, does not match up with our nature–nurture intuitions. We want to know how "important" the roles of genes and environment are to the development of a trait, but in focusing on "important" maybe we're emphasizing the wrong thing. First of all, genes and environment are both crucial to *every* trait; without genes the environment would have nothing to work on, and too, genes cannot develop in a vacuum. Even more important, because nature–nurture questions look at the differences among people, the cause of a given trait depends not only on the trait itself, but also on the differences in that trait between members of the group being studied.

The classic example of the heritability coefficient defying intuition is the trait of having two arms. No one would argue against the development of arms being a biological, genetic process. But fraternal twins are just as similar for "two-armedness" as identical twins, resulting in a heritability coefficient of zero for the trait of having two arms. Normally, according to the heritability model, this result (coefficient of zero) would suggest all nurture, no nature, but we know that's not the case. The reason this result is not a tip-off that arm development is less genetic than we imagine is because people *do not vary* in the genes related to arm development—which essentially upends the heritability formula. In fact, in this instance, the opposite is likely true: the extent that people differ in arm number is likely the result of accidents and, therefore, environmental. For reasons like these, we always have to be very careful when asking nature–nurture questions, especially when we try to express the answer in terms of a single number. The heritability of a trait is not simply a property of that trait, but a property of the trait in a particular context of relevant genes and environmental factors.

Another issue with the heritability coefficient is that it divides traits' determinants into two portions—genes and environment—which are then calculated together for the total variability. This is a little like asking how much of the experience of a symphony comes from the horns and how much from the strings; the ways instruments or genes integrate is more complex than that. It turns out to be the case that, for many traits, genetic differences affect behavior under some environmental circumstances but not others—a phenomenon called gene-environment interaction, or G x E. In one well-known example, Caspi et al. (2002) showed that among maltreated children, those who carried a particular allele of the MAOA gene showed a predisposition to violence and antisocial behavior, while those with other alleles did not. Whereas, in children who had not been maltreated, the gene had no effect. Making matters even more complicated are very recent studies of what is known as epigenetics (see module, "Epigenetics" http://noba.to/37p5cb8v), a process in which the DNA itself is modified by environmental events, and those genetic changes transmitted to children.

The answer to the nature –nurture question has not turned out to be as straightforward as we would like. The many questions we can ask about the relationships among genes, environments, and human traits may have many different answers, and the answer to one tells us little about the answers to the others. [Image: Sundaram Ramaswamy, https://goo.gl/Bv8lp6, CC BY 2.0, https://goo.gl/9uSnqN]

Some common questions about nature–nurture are, how susceptible is a trait to change, how malleable is it, and do we "have a choice" about it? These questions are much more complex than they may seem at first glance. For example, phenylketonuria is an inborn error of metabolism caused by a single gene; it prevents the body from metabolizing phenylalanine. Untreated, it causes intellectual disability and death. But it can be treated effectively by a straightforward environmental intervention: avoiding foods containing phenylalanine. Height seems like a trait firmly rooted in our nature and unchangeable, but the average height of many populations in Asia and Europe has increased significantly in the past 100 years, due to changes in diet and the alleviation of poverty. Even the most modern genetics has not provided definitive answers to nature–nurture questions. When it was first becoming possible to measure the DNA sequences of individual people, it was widely thought that we would quickly progress to finding the specific genes that account for behavioral characteristics, but that hasn't happened. There are a few rare genes that have been found to have significant (almost always negative) effects, such as the single gene that causes Huntington's disease, or the Apolipoprotein gene that causes early onset dementia in a small percentage of Alzheimer's cases. Aside from these rare genes of great effect, however, the genetic impact on behavior is broken up over many genes, each with very small effects. For most behavioral traits, the effects are so small and distributed across so many genes that we have not been able to catalog them in a meaningful way. In fact, the same is true of environmental effects. We know that extreme environmental hardship causes catastrophic effects for many behavioral outcomes, but fortunately extreme environmental hardship is very rare. Within the normal range of environmental events, those responsible for differences (e.g., why some children in a suburban third-grade classroom perform better than others) are much more difficult to grasp.

The difficulties with finding clear-cut solutions to nature–nurture problems bring us back to the other great questions about our relationship with the natural world: the mind-body problem and free will. Investigations into what we mean when we say we are aware of something reveal that consciousness is not simply the product of a particular area of the brain, nor does choice turn out to be an orderly activity that we can apply to some behaviors but not others. So it is with nature and nurture: What at first may seem to be a straightforward matter, able to be indexed with a single number, becomes more and more complicated the closer we look. The many questions we can ask about the intersection among genes, environments, and human traits—how sensitive are traits to environmental change, and how common are those influential environments; are parents or culture more relevant; how sensitive are traits to differences in genes, and how much do the relevant genes vary in a particular population; does the trait involve a single gene or a great many genes; is the trait more easily described in genetic or more-complex behavioral terms?—may have different answers, and the answer to one tells us little about the answers to the others.

It is tempting to predict that the more we understand the wide-ranging effects of genetic differences on all human characteristics—especially behavioral ones—our cultural, ethical, legal, and personal ways of thinking about ourselves will have to undergo profound changes in response. Perhaps criminal proceedings will consider genetic background. Parents, presented with the genetic sequence of their children, will be faced with difficult decisions about reproduction. These hopes or fears are often exaggerated. In some ways, our thinking may need to change—for example, when we consider the meaning behind the fundamental American principle that all men are created equal. Human beings differ, and like all evolved organisms they differ genetically. The Declaration of Independence predates Darwin and Mendel, but it is hard to imagine that Jefferson—whose genius encompassed botany as well as moral philosophy—would have been alarmed to learn about the genetic diversity of organisms. One of the most important things modern genetics has taught us is that almost all human behavior is too complex to be nailed down, even from the most complete genetic information, unless we're looking at identical twins. The science of nature and nurture has demonstrated that genetic differences among people are vital to human moral equality, freedom, and self-determination, not opposed to them. As Mordecai Kaplan said about the role of the past in Jewish theology, genetics gets a vote, not a veto, in the determination of human behavior. We should indulge our fascination with nature–nurture while resisting the temptation to oversimplify it.

Outside Resources

Web: Institute for Behavioral Genetics
http://www.colorado.edu/ibg/

Discussion Questions

1. Is your personality more like one of your parents than the other? If you have a sibling, is his or her personality like yours? In your family, how did these similarities and differences develop? What do you think caused them?
2. Can you think of a human characteristic for which genetic differences would play almost no role? Defend your choice.
3. Do you think the time will come when we will be able to predict almost everything about someone by examining their DNA on the day they are born?
4. Identical twins are more similar than fraternal twins for the trait of aggressiveness, as well as for criminal behavior. Do these facts have implications for the courtroom? If it can be shown that a violent criminal had violent parents, should it make a difference in culpability or sentencing?

Vocabulary

Adoption study
A behavior genetic research method that involves comparison of adopted children to their adoptive and biological parents.

Behavioral genetics
The empirical science of how genes and environments combine to generate behavior.

Heritability coefficient
An easily misinterpreted statistical construct that purports to measure the role of genetics in the explanation of differences among individuals.

Quantitative genetics
Scientific and mathematical methods for inferring genetic and environmental processes based on the degree of genetic and environmental similarity among organisms.

Twin studies
A behavior genetic research method that involves comparison of the similarity of identical (monozygotic; MZ) and fraternal (dizygotic; DZ) twins.

References

Bouchard, T. J., Lykken, D. T., McGue, M., & Segal, N. L. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. Science, 250(4978), 223–228.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A. & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. Science, 297(5582), 851–854.

McGue, M., & Lykken, D. T. (1992). Genetic influence on risk of divorce. Psychological Science, 3(6), 368–373.

Plomin, R., Corley, R., DeFries, J. C., & Fulker, D. W. (1990). Individual differences in television viewing in early childhood: Nature as well as nurture. Psychological Science, 1(6), 371–377.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2012). Behavioral genetics. New York, NY: Worth Publishers.

Scott, J. P., & Fuller, J. L. (1998). Genetics and the social behavior of the dog. Chicago, IL: University of Chicago Press.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. Current Directions in Psychological Science, 9(5), 160–164.

Sensation and Perception

12
Vision

Simona Buetti & Alejandro Lleras

Vision is the sensory modality that transforms light into a psychological experience of the world around you, with minimal bodily effort. This module provides an overview of the most significant steps in this transformation and strategies that your brain uses to achieve this visual understanding of the environment.

Learning Objectives

- Describe how the eye transforms light information into neural energy.
- Understand what sorts of information the brain is interested in extracting from the environment and why it is useful.
- Describe how the visual system has adapted to deal with different lighting conditions.
- Understand the value of having two eyes.
- Understand why we have color vision.
- Understand the interdependence between vision and other brain functions.

What Is Vision?

Think about the spectacle of a starry night. You look up at the sky, and thousands of photons from distant stars come crashing into your retina, a light-sensitive structure at the back of your eyeball. These photons are millions of years old and have survived a trip across the universe, only to run into one of your photoreceptors. Tough luck: in one thousandth of a second, this little bit of light energy becomes the fuel to a photochemical reaction known as photoactivation. The light energy becomes neural energy and triggers a cascade of neural activity that, a few hundredths of a second later, will result in your becoming aware of that distant star. You and the universe united by photons. That is the amazing power of vision. Light brings the world to you. Without moving, you know what's out there. You can recognize friends coming to meet you before you are able to hear them coming, ripe fruits from green ones on trees without having to taste them and before reaching out to grab them. You can also tell how quickly a ball is moving in your direction (Will it hit you? Can you hit it?).

How does all of that happen? First, light enters the eyeball through a tiny hole known as the pupil and, thanks to the refractive properties of your cornea and lens, this light signal gets projected sharply into the retina (see Outside Resources for links to a more detailed description of the eye structure). There, light is transduced into neural energy by about 200 million photoreceptor cells.

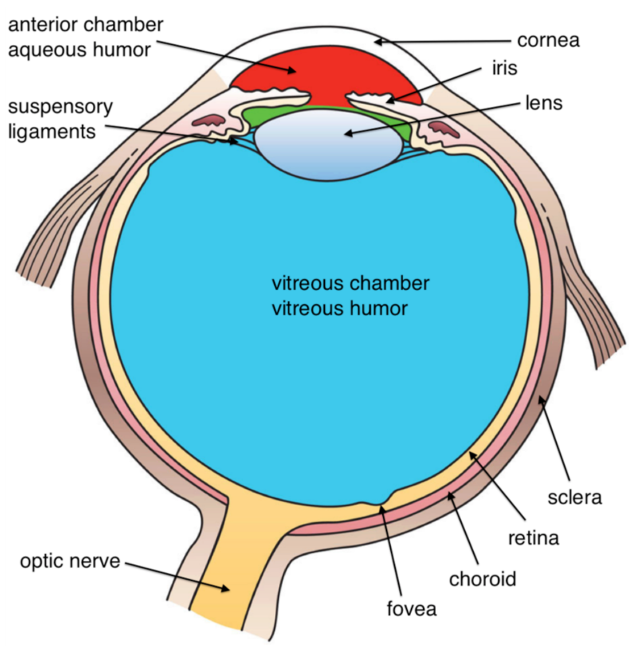


Diagram of the human eye [Image: Holly Fischer, https://goo.gl/cMOQdh, CC BY 3.0, https://goo.gl/EHHVdU]

This is where the information carried by the light about distant objects and colors starts being encoded by our brain. There are two different types of photoreceptors: rods and cones. The human eye contains more rods than cones. Rods give us sensitivity under dim lighting conditions and allow us to see at night. Cones allow us to see fine details in bright light and give us the sensation of color. Cones are tightly packed around the fovea (the central region of the retina behind your pupil) and more sparsely elsewhere. Rods populate the periphery (the region surrounding the fovea) and are almost absent from the fovea.

But vision is far more complex than just catching photons. The information encoded by the photoreceptors undergoes a rapid and continuous set of ever more complex analysis so that, eventually, you can make sense of what's out there. At the fovea, visual information is encoded separately from tiny portions of the world (each about half the width of a human hair viewed at arm's length) so that eventually the brain can reconstruct in great detail fine visual differences from locations at which you are directly looking. This fine level of encoding requires lots of light and it is slow going (neurally speaking). In contrast, in the periphery, there is a different encoding strategy: detail is sacrificed in exchange for sensitivity. Information is summed across larger sections of the world. This aggregation occurs quickly and allows you to detect dim signals under very low levels of light, as well as detect sudden movements in your peripheral vision.

The Importance of Contrast

What happens next? Well, you might think that the eye would do something like record the amount of light at each location in the world and then send this information to the visual-processing areas of the brain (an astounding 30% of the cortex is influenced by visual signals!). But, in fact, that is not what eyes do. As soon as photoreceptors capture light, the nervous system gets busy analyzing *differences* in light, and it is these differences that get transmitted to the brain. The brain, it turns out, cares little about the overall amount of light coming from a specific part of the world, or in the scene overall. Rather, it wants to know: does the light coming from this one point differ from the light coming from the point next to it? Place your hand on the table in front of you. The contour of your hand is actually determined by the difference in light—the **contrast**—between the light coming from the skin in your hand and the light coming from the table underneath. To find the contour of your hand, we simply need to find the regions in the image where the difference in light between two adjacent points is maximal. Two points on your skin will reflect similar levels of light back to you, as will two points on the table. On the other hand, two points that fall on either side of the boundary contour between your hand and the table will reflect very different light.

The fact that the brain is interested in coding contrast in the world reveals something deeply important about the forces that drove the evolution of our brain: encoding the absolute amount of light in the world tells us little about what is out there. But if your brain can detect the sudden appearance of a *difference* in light somewhere in front of you, then it must be that something new is there. That contrast signal is *information*. That information may represent something that you like (food, a friend) or something dangerous approaching (a tiger, a cliff). The rest of your visual system will work hard to determine what that thing is, but as quickly as 10ms after light enters your eyes, ganglion cells in your retinae have already encoded all the differences in light from the world in front of you.

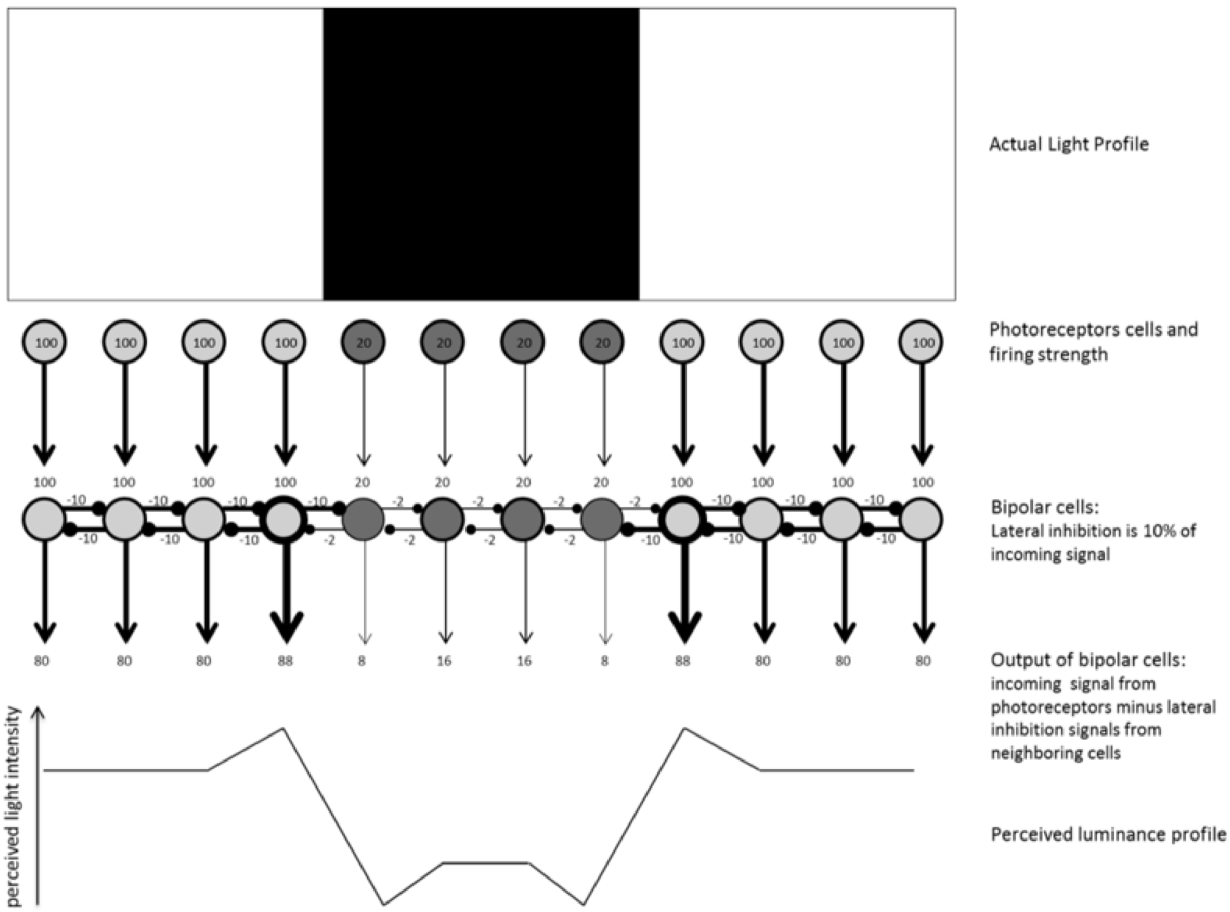


Figure 1. Illustration of Lateral Inhibition at work. The top of the figure shows a black stripe on a white background. The first row of circles illustrates photoreceptors responding in a graded fashion: the more light hits them, the more they fire. The numbers inside the circles represent how much these cells are firing, and the thickness of lines is also meant to illustrate the strength of neural firing. These photoreceptors activate the next layer of neurons in the retina: bipolar cells. These cells produce lateral inhibition signals, depicted by the horizontal lines that end with a small circle. The inhibition signals are proportional (here, 10% for ease) to the excitatory input they receive. Cells receiving 100 units will inhibit their neighbors by 10 units. Cells receiving 20 units will inhibit their neighbors by 2 units. The output of a bipolar cell will be determined by the input it receives minus all the lateral inhibition signals from its neighbors. As a result of the inhibition, notice how on the bright side of the edges, the firing rates are the highest (88) compared to nearby neurons just coding bright light (80). These higher values near the edge occur because these cells receive a comparatively small amount of inhibition from their "dark-side" neighbor (-2). Similarly, on the dark side of the edge, the firing rates are the lowest (8) of all the dark region (16) because these cells receive a comparatively large amount of inhibition from their "bright-side" neighbor (-10). Overall, the image is coded as a black stripe surrounded by brighter light, but now, thanks to lateral inhibition, all the edges in the image have been emphasized (enhanced), as illustrated by the perceived luminance profile at the bottom of the image.

Contrast is so important that your neurons go out of their way not only to encode differences in light but to exaggerate those differences for you, lest you miss them. Neurons achieve this via a process known as lateral inhibition. When a neuron is firing in response to light, it produces two signals: an output signal to pass on to the next level in vision, and a lateral signal to inhibit all neurons that are next to it. This makes sense on the assumption that nearby neurons are likely responding to the same light coming from nearby locations, so this information is somewhat redundant. The magnitude of the lateral inhibitory signal a neuron produces is proportional to the excitatory input that neuron receives: the more a neuron fires, the stronger the inhibition it produces. Figure 1 illustrates how lateral inhibition amplifies contrast signals at the edges of surfaces.

Sensitivity to Different Light Conditions

Rods and cones work in tandem to help you adjust when moving between extremes of dark and light. [Image: Pexels, CC0 Public Domain, https://goo.gl/m25gce]

Let's think for a moment about the range of conditions in which your visual system must operate day in and day out. When you take a walk outdoors on a sunny day, as many as billions of photons enter your eyeballs every second. In contrast, when you wake up in the middle of the night in a dark room, there might be as little as a few hundred photons per second entering your eyes. To deal with these extremes, the visual system relies on the different properties of the two types of photoreceptors. Rods are mostly responsible for processing light when photons are scarce (just a single photon can make a rod fire!), but it takes time to replenish the visual pigment that rods require for photoactivation. So, under bright conditions, rods are quickly bleached (Stuart & Brige, 1996) and cannot keep up with the constant barrage of photons hitting them. That's when the cones become useful. Cones require more photons to fire and, more critically, their photopigments replenish much faster than rods' photopigments, allowing them to keep up when photons are abundant.

What happens when you abruptly change lighting conditions? Under bright light, your rods are bleached. When you move into a dark environment, it will take time (up to 30 minutes) before they chemically recover (Hurley, 2002). Once they do, you will begin to see things around you that initially you could not. This phenomenon is called **dark adaptation**. When you go from dark to bright light (as you exit a tunnel on a highway, for instance), your rods will be bleached in a blaze and you will be blinded by the sudden light for about 1 second. However, your cones are ready to fire! Their firing will take over and you will quickly begin to see at this higher level of light.

A similar, but more subtle, adjustment occurs when the change in lighting is not so drastic. Think about your experience of reading a book at night in your bed compared to reading outdoors: the room may feel to you fairly well illuminated (enough so you can read) but the light bulbs in your room are not producing the billions of photons that you encounter outside. In both cases, you feel that your experience is that of a well-lit environment. You don't feel one experience as millions of times brighter than the other. This is because vision (as much of perception) is not proportional: seeing twice as many photons does not produce a sensation of seeing twice as bright a light. The visual system tunes into the current experience by favoring a range of contrast values that is most informative in that environment (Gardner et al., 2005). This is the concept of **contrast gain**: the visual system determines the mean contrast in a scene and represents values around that mean contrast best, while ignoring smaller contrast differences. (See the Outside Resources section for a demonstration.)

The Reconstruction Process

What happens once information leaves your eyes and enters the brain? Neurons project first into the thalamus, in a section known as the lateral geniculate nucleus. The information then splits and projects towards two different parts of the brain. Most of the computations regarding reflexive eye movements are computed in subcortical regions, the evolutionarily older part of the brain. Reflexive eye movements allow you to quickly orient your eyes towards areas of interest and to track objects as they move. The more complex computations, those that eventually allow you to have a visual experience of the world, all happen in the cortex, the evolutionarily newer region of the brain. The first stop in the cortex is at the **primary visual cortex** (also known as V1). Here, the "reconstruction" process begins in earnest: based on the contrast information arriving from the eyes, neurons will start computing information about color and simple lines, detecting various orientations and thicknesses. Small-scale motion signals are also computed (Hubel & Wiesel, 1962).

As information begins to flow towards other "higher" areas of the system, more complex computations are performed. For example, edges are assigned to the object to which they belong, backgrounds are separated from foregrounds, colors are assigned to surfaces, and the global motion of objects is computed. Many of these computations occur in specialized brain areas. For instance, an area called MT processes global-motion information; the parahippocampal place area identifies locations and scenes; the fusiform face area specializes in identifying objects for which fine discriminations are required, like faces. There is even a brain region specialized in letter and word processing. These visual-recognition areas are located along the *ventral pathway* of the brain (also known as the What pathway). Other brain regions along the *dorsal pathway* (or Where-and-How pathway) will compute information about self- and object-motion, allowing you to interact with objects, navigate the environment, and avoid obstacles (Goodale and Milner, 1992).

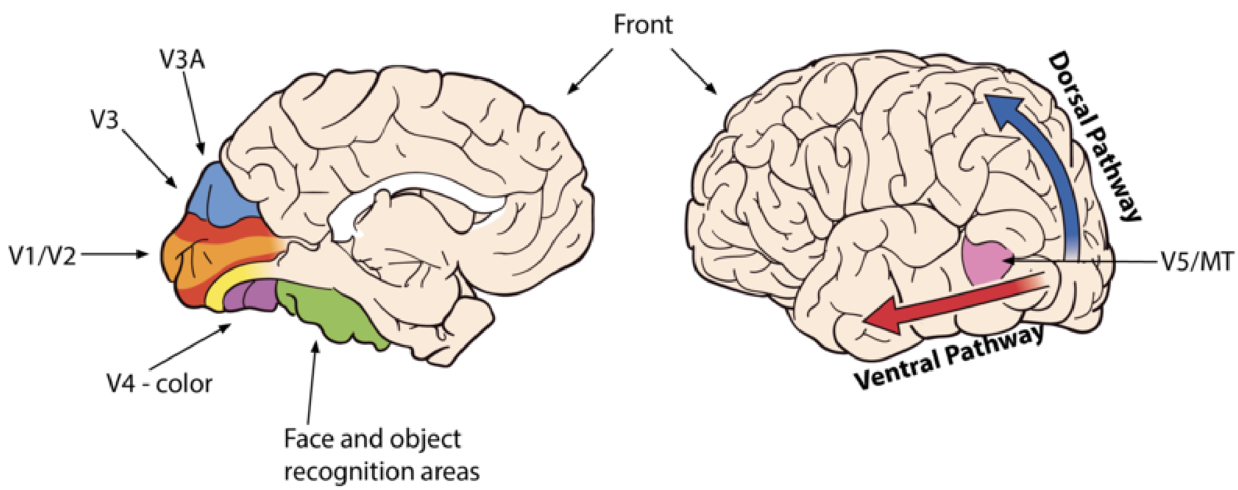


Figure 2: Areas of the brain

Now that you have a basic understanding of how your visual system works, you can ask yourself the question: why do you have two eyes? Everything that we discussed so far could be computed with information coming from a single eye. So why two? Looking at the animal kingdom gives us a clue. Animals who tend to be prey have eyes located on opposite sides of their skull. This allows them to detect predators whenever one appears anywhere around them. Humans, like most predators, have two eyes pointing in the same direction, encoding almost the exact scene twice. This redundancy gives us a binocular advantage: having two eyes not only provides you with two chances at catching a signal in front of you, but the minute difference in perspective that you get from each eye is used by your brain to reconstruct the sense of three-dimensional space. You can get an estimate of how far distant objects are from you, their size, and their volume. This is no easy feat: the signal in each eye is a two-dimensional projection of the world, like two separate pictures drawn upon your retinae. Yet, your brain effortlessly provides you with a sense of depth by combining those two signals. This 3-D reconstruction process also relies heavily on all the knowledge you acquired through experience about spatial information. For instance, your visual system learns to interpret how the volume, distance, and size of objects change as they move closer or farther from you. (See the Outside Resources section for demonstrations.)

The Experience of Color

Perhaps one of the most beautiful aspects of vision is the richness of the color experience that it provides us. One of the challenges that we have as scientists is to understand why the human color experience is what it is. Perhaps you have heard that dogs only have 2 types of color photoreceptors, whereas humans have 3, chickens have 4, and mantis shrimp have 16. Why is there such variation across species? Scientists believe each species has evolved with different needs and uses color perception to signal information about food, reproduction, and health that are unique to their species. For example, humans have a specific sensitivity that allows you to detect slight changes in skin tone. You can tell when someone is embarrassed, aroused, or ill. Detecting these subtle signals is adaptive in a social species like ours.

How is color coded in the brain? The two leading theories of color perception were proposed in the mid-19th century, about 100 years before physiological evidence was found to corroborate them both (Svaetichin, 1956). **Trichromacy theory**, proposed by Young (1802) and Helmholtz (1867), proposed that the eye had three different types of color-sensitive cells based on the observation that any one color can be reproduced by combining lights from three lamps of different hue. If you can adjust separately the intensity of each light, at some point you will find the right combination of the three lights to match any color in the world. This principle is used today on TVs, computer screens, and any colored display. If you look closely enough at a pixel, you will find that it is composed of a blue, a red, and a green light, of varying intensities. Regarding the retina, humans have three types of cones: S-cones, M-cones, and L-cones (also known as blue, green, and red cones, respectively) that are sensitive to three different wavelengths of light.

Around the same time, Hering made a puzzling discovery: some colors are impossible to create. Whereas you can make yellowish greens, bluish reds, greenish blues, and reddish yellows by combining two colors, you can never make a reddish green or a bluish yellow. This observation led Hering (1892) to propose the **Opponent Process theory of color**: color is coded via three opponent channels (red-green, blue-yellow, and black-white). Within each channel, a comparison is constantly computed between the two elements in the pair. In other words, colors are encoded as *differences* between two hues and not as simple combinations of hues. Again, what matters to the brain is *contrast*. When one element is stronger than the other, the stronger color is perceived and the weaker one is suppressed. You can experience this phenomenon by following the link below.

http://nobaproject.com/assets/modules/module-visio...

When both colors in a pair are present to equal extents, the color perception is canceled and we perceive a level of grey. This is why you cannot see a reddish green or a bluish yellow: they cancel each other out. By the way, if you are wondering where the yellow signal comes from, it turns out that it is computed by averaging the M- and L-cone signals. Are these colors uniquely human colors? Some think that they are: the red-green contrast, for example, is finely tuned to detect changes in human skin tone so you can tell when someone blushes or becomes pale. So, the next time you go out for a walk with your dog, look at the sunset and ask yourself, what color does my dog see? Probably none of the orange hues you do!

So now, you can ask yourself the question: do all humans experience color in the same way? Color-blind people, as you can imagine, do not see all the colors that the rest of us see, and this is due to the fact that they lack one (or more) cones in their retina. Incidentally, there are a few women who actually have four different sets of cones in their eyes, and recent research suggests that their experience of color can be (but not always is) richer than the one from three-coned people. A slightly different question, though, is whether all three-coned people have the same internal experiences of colors: is the red inside your head the same red inside your mom’s head? That is an almost impossible question to answer that has been debated by philosophers for millennia, yet recent data suggests that there might in fact be cultural differences in the way we perceive color. As it turns out, not all cultures categorize colors in the same way, for example. And some groups “see” different shades of what we in the Western world would call the “same” color, as categorically different colors. The Berinmo tribe in New Guinea, for instance, appear to experience green shades that denote leaves that are alive as belonging to an entirely different color category than the sort of green shades that denote dying leaves. Russians, too, appear to experience light and dark shades of blue as different categories of colors, in a way that most Westerners do not. Further, current brain imaging research suggests that people’s brains change (increase in white-matter volume) when they learn new color categories! These are intriguing and suggestive findings, for certain, that seem to indicate that our cultural environment may in fact have some (small) but definite impact on how people use and experience colors across the globe.

Integration with Other Modalities

Vision is not an encapsulated system. It interacts with and depends on other sensory modalities. For example, when you move your head in one direction, your eyes reflexively move in the opposite direction to compensate, allowing you to maintain your gaze on the object that you are looking at. This reflex is called the **vestibulo-ocular reflex**. It is achieved by integrating information from both the visual and the vestibular system (which knows about body motion and position). You can experience this compensation quite simply. First, while you keep your head still and your gaze looking straight ahead, wave your finger in front of you from side to side. Notice how the image of the finger appears blurry. Now, keep your finger steady and look at it while you move your head from side to side. Notice how your eyes reflexively move to compensate the movement of your head and how the image of the finger stays sharp and stable. Vision also interacts with your proprioceptive system, to help you find where all your body parts are, and with your auditory system, to help you understand the sounds people make when they speak. You can learn more about this in the Noba module about multimodal perception (**http://noba.to/cezw4qyn**).

Finally, vision is also often implicated in a blending-of-sensations phenomenon known as **synesthesia**. Synesthesia occurs when one sensory signal gives rise to two or more sensations. The most common type is *grapheme-color* synesthesia. About 1 in 200 individuals experience a sensation of color associated with specific letters, numbers, or words: the number 1 might always be seen as red, the number 2 as orange, etc. But the more fascinating forms of synesthesia blend sensations from entirely different sensory modalities, like taste and color or music and color: the taste of chicken might elicit a sensation of green, for example, and the timbre of violin a deep purple.

Concluding Remarks

We are at an exciting moment in our scientific understanding of vision. We have just begun to get a functional understanding of the visual system. It is not sufficiently evolved for us to recreate artificial visual systems (i.e., we still cannot make robots that "see" and understand light signals as we do), but we are getting there. Just recently, major breakthroughs in vision science have allowed researchers to significantly improve retinal prosthetics: photosensitive circuits that can be implanted on the back of the eyeball of blind people that connect to visual areas of the brain and have the capacity to partially restore a "visual experience" to these patients (Nirenberg & Pandarinath, 2012). And using functional magnetic brain imaging, we can now "decode" from your brain activity the images that you saw in your dreams while you were asleep (Horikawa, Tamaki, Miyawaki, & Kamitani, 2013)! Yet, there is still so much more to understand. Consider this: if vision is a construction process that takes time, whatever we see *now* is no longer what is front of us. Yet, humans can do amazing time-sensitive feats like hitting a 90-mph fastball in a baseball game. It appears then that a fundamental function of vision is not just to know what is happening around you now, but actually to make an accurate inference about what you are about to see *next* (Enns & Lleras, 2008), so that you can keep up with the world. Understanding how this future-oriented, *predictive* function of vision is achieved in the brain is probably the next big challenge in this fascinating realm of research.

Outside Resources

Video: Acquired knowledge and its impact on our three-dimensional interpretation of the world - 3D Street Art
https://youtu.be/bkJh3xxC6UU

Video: Acquired knowledge and its impact on our three-dimensional interpretation of the world - Anamorphic Illusions
http://www.youtube.com/watch?v=tBNHPk-Lnkk

Video: Acquired knowledge and its impact on our three-dimensional interpretation of the world - Optical Illusion
http://www.youtube.com/watch?v=YjmHofJ2da0&feature=related

Web: Amazing library with visual phenomena and optical illusions, explained
http://michaelbach.de/ot/index.html

Web: Anatomy of the eye
http://www.eyecareamerica.org/eyecare/anatomy/

Web: Demonstration of contrast gain adaptation
http://www.michaelbach.de/ot/lum_contrast-adapt/

Web: Demonstration of illusory contours and lateral inhibition. Mach bands
http://michaelbach.de/ot/lum-MachBands/index.html

Web: Demonstration of illusory contrast and lateral inhibition. The Hermann grid
http://michaelbach.de/ot/lum_herGrid/

Web: Further information regarding what and where/how pathways
http://www.scholarpedia.org/article/What_and_where_pathways

Discussion Questions

1. When running in the dark, it is recommended that you never look straight at the ground. Why? What would be a better strategy to avoid obstacles?

2. The majority of ganglion cells in the eye specialize in detecting drops in the amount of light coming from a given location. That is, they increase their firing rate when they detect less light coming from a specific location. Why might the absence of light be more important than the presence of light? Why would it be evolutionarily advantageous to code this type of information?
3. There is a hole in each one of your eyeballs called the optic disk. This is where veins enter the eyeball and where neurons (the axons of the ganglion cells) exit the eyeball. Why do you not see two holes in the world all the time? Close one eye now. Why do you not see a hole in the world now? To "experience" a blind spot, follow the instructions in this website: http://michaelbach.de/ot/cog_blindSpot/index.html
4. Imagine you were given the task of testing the color-perception abilities of a newly discovered species of monkeys in the South Pacific. How would you go about it?
5. An important aspect of emotions is that we sense them in ourselves much in the same way as we sense other perceptions like vision. Can you think of an example where the concept of contrast gain can be used to understand people's responses to emotional events?

Vocabulary

Binocular advantage
Benefits from having two eyes as opposed to a single eye.

Cones
Photoreceptors that operate in lighted environments and can encode fine visual details. There are three different kinds (S or blue, M or green and L or red) that are each sensitive to slightly different types of light. Combined, these three types of cones allow you to have color vision.

Contrast
Relative difference in the amount and type of light coming from two nearby locations.

Contrast gain
Process where the sensitivity of your visual system can be tuned to be most sensitive to the levels of contrast that are most prevalent in the environment.

Dark adaptation
Process that allows you to become sensitive to very small levels of light, so that you can actually see in the near-absence of light.

Lateral inhibition
A signal produced by a neuron aimed at suppressing the response of nearby neurons.

Opponent Process Theory
Theory of color vision that assumes there are four different basic colors, organized into two pairs (red/green and blue/yellow) and proposes that colors in the world are encoded in terms of the opponency (or difference) between the colors in each pair. There is an additional black/white pair responsible for coding light contrast.

Photoactivation
A photochemical reaction that occurs when light hits photoreceptors, producing a neural signal.

Primary visual cortex (V1)
Brain region located in the occipital cortex (toward the back of the head) responsible for processing basic visual information like the detection, thickness, and orientation of simple lines, color, and small-scale motion.

Rods
Photoreceptors that are very sensitive to light and are mostly responsible for night vision.

Synesthesia
The blending of two or more sensory experiences, or the automatic activation of a secondary (indirect) sensory experience due to certain aspects of the primary (direct) sensory stimulation.

Trichromacy theory
Theory that proposes that all of your color perception is fundamentally based on the combination of three (not two, not four) different color signals.

Vestibulo-ocular reflex
Coordination of motion information with visual information that allows you to maintain your gaze on an object while you move.

What pathway
Pathway of neural processing in the brain that is responsible for your ability to recognize what is around you.

Where-and-How pathway
Pathway of neural processing in the brain that is responsible for you knowing where things are in the world and how to interact with them.

References

Enns, J. T., & Lleras, A. (2008). New evidence for prediction in human vision. *Trends in Cognitive Sciences, 12*¸ 327–333.

Gardner, J. L., Sun, P., Waggoner, R. A. , Ueno, K., Tanaka, K., & Cheng, K. (2005). Contrast adaptation and representation in human early visual cortex. *Neuron, 47*, 607–620.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience, 15*, 20–25.

Helmholtz, H. von. (1867). *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss.

Hering, E. (1892). *Grundzu ge der Lehre vom Lichtsinn*. Berlin, Germany: Springer.

Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science, 340*(6132), 639–642.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106–154.

Hurley, J. B. (2002). Shedding light on adaptation. *Journal of General Physiology, 119*, 125–128.

Nirenberg, S., & Pandarinath, C. (2012). Retinal prosthetic strategy with the capacity to restore normal vision. *Proceedings of the National Academy of Sciences, 109* (37), 15012–15017.

Stuart, J. A., & Brige, R. R. (1996). Characterization of the primary photochemical events in bacteriorhodopsin and rhodopsin. In A. G. Lee (Ed.), *Rhodopsin and G-protein linked receptors* (Part A, Vol. 2, pp. 33–140). Greenwich, CT: JAI.

Svaetichin, G. (1956). Spectral response curves from single cones, *Actaphysiologica Scandinavia, Suppl. 134*, 17–46.

Young, T. (1802). Bakerian lecture: On the theory of light and colours. *Philosophical Transaction of the Royal Society London, 92*, 12–48.

13
Hearing

Andrew J. Oxenham

Hearing allows us to perceive the world of acoustic vibrations all around us, and provides us with our most important channels of communication. This module reviews the basic mechanisms of hearing, beginning with the anatomy and physiology of the ear and a brief review of the auditory pathways up to the auditory cortex. An outline of the basic perceptual attributes of sound, including loudness, pitch, and timbre, is followed by a review of the principles of tonotopic organization, established in the cochlea. An overview of masking and frequency selectivity is followed by a review of the perception and neural mechanisms underlying spatial hearing. Finally, an overview is provided of auditory scene analysis, which tackles the important question of how the auditory system is able to make sense of the complex mixtures of sounds that are encountered in everyday acoustic environments.

Learning Objectives

- Describe the basic auditory attributes of sound.
- Describe the structure and general function of the auditory pathways from the outer ear to the auditory cortex.
- Discuss ways in which we are able to locate sounds in space.
- Describe various acoustic cues that contribute to our ability to perceptually segregate simultaneously arriving sounds.

Introduction

Hearing forms a crucial part of our everyday life. Most of our communication with others, via speech or music, reaches us through the ears. Indeed, a saying, often attributed to Helen Keller, is that blindness separates us from things, but deafness separates us from people. The ears respond to acoustic information, or sound —tiny and rapid variations in air pressure. Sound waves travel from the source and produce pressure variations in the listener's ear canals, causing the eardrums (or tympanic membranes) to vibrate. This module provides an overview of the events that follow, which convert these simple mechanical vibrations into our rich experience known as hearing, or auditory perception.

Hearing provides us with our most important connection to the people around us. [Image: Bindaas Madhavi, https://goo.gl/Sv6TtR, CC BY-NC-ND 2.0, https://goo.gl/62XJAl]

Perceptual Attributes of Sound

There are many ways to describe a sound, but the perceptual attributes of a sound can typically be divided into three main categories—namely, loudness, pitch, and timbre. Although all three refer to perception, and not to the physical sounds themselves, they are strongly related to various physical variables.

Loudness

The most direct physical correlate of loudness is sound intensity (or sound pressure) measured close to the eardrum. However, many other factors also influence the loudness of a sound, including its frequency content, its duration, and the context in which it is presented. Some of the earliest psychophysical studies of auditory perception, going back more than a century, were aimed at examining the relationships between perceived loudness, the physical sound intensity, and the just-noticeable differences in loudness (Fechner, 1860; Stevens, 1957). A great deal of time and effort has been spent refining various measurement methods. These methods involve techniques such as magnitude estimation, where a series of sounds (often sinusoids, or pure tones of single frequency) are presented sequentially at different sound levels, and subjects are asked to assign numbers to each tone, corresponding to the perceived loudness. Other studies have examined how loudness changes as a function of the frequency of a tone, resulting in the international standard iso-loudness-level contours (ISO, 2003), which are used in many areas of industry to assess noise and annoyance issues. Such studies have led to the development of computational models that are designed to predict the loudness of arbitrary sounds (e.g., Moore, Glasberg, & Baer, 1997).

Pitch

Pitch is crucial to our perception and understanding of music and language. [Image: xroper7, https://goo.gl/1E4sJY, CC BY-NC 2.0, https://goo.gl/tgFydH]

Pitch plays a crucial role in acoustic communication. Pitch variations over time provide the basis of melody for most types of music; pitch contours in speech provide us with important prosodic information in non-tone languages, such as English, and help define the meaning of words in tone languages, such as Mandarin Chinese. Pitch is essentially the perceptual correlate of waveform periodicity, or repetition rate: The faster a waveform repeats over time, the higher is its perceived pitch. The most common pitch-evoking sounds are known as harmonic complex tones. They are complex because they consist of more than one frequency, and they are harmonic because the frequencies are all integer multiples of a common fundamental frequency (F0). For instance, a harmonic complex tone with a F0 of 100 Hz would also contain energy at frequencies of 200, 300, 400 Hz, and so on. These higher frequencies are known as harmonics or overtones, and they also play an important role in determining the pitch of a sound. In fact, even if the energy at the F0 is absent or masked, we generally still perceive the remaining sound to have a pitch corresponding to the F0. This phenomenon is known as the "pitch of the missing fundamental," and it has played an important role in the formation of theories and models about pitch (de Cheveigné, 2005). We hear pitch with sufficient accuracy to perceive melodies over a range of F0s from about 30 Hz (Pressnitzer, Patterson, & Krumbholz, 2001) up to about 4–5 kHz (Attneave & Olson, 1971; Oxenham, Micheyl, Keebler, Loper, & Santurette, 2011). This range also corresponds quite well to the range covered by musical instruments; for instance, the modern grand piano has notes that extend from 27.5 Hz to 4,186 Hz. We are able to discriminate changes in frequency above 5,000 Hz, but we are no longer very accurate in recognizing melodies or judging musical intervals.

Timbre

Timbre refers to the quality of sound, and is often described using words such as bright, dull, harsh, and hollow. Technically, timbre includes anything that allows us to distinguish two sounds that have the same loudness, pitch, and duration. For instance, a violin and a piano playing the same note sound very different, based on their sound quality or timbre.

An important aspect of timbre is the spectral content of a sound. Sounds with more high-frequency energy tend to sound brighter, tinnier, or harsher than sounds with more low-frequency content, which might be described as deep, rich, or dull. Other important aspects of timbre include the temporal envelope (or outline) of the sound, especially how it begins and ends. For instance, a piano has a rapid onset, or attack, produced by the hammer striking the string, whereas the attack of a clarinet note can be much more gradual. Artificially changing the onset of a piano note by, for instance, playing a recording backwards, can dramatically alter its character so that it is no longer recognizable as a piano note. In general, the overall spectral content and the temporal envelope can provide a good first approximation to any sound, but it turns out that subtle changes in the spectrum over time (or spectro-temporal variations) are crucial in creating plausible imitations of natural musical instruments (Risset & Wessel, 1999).

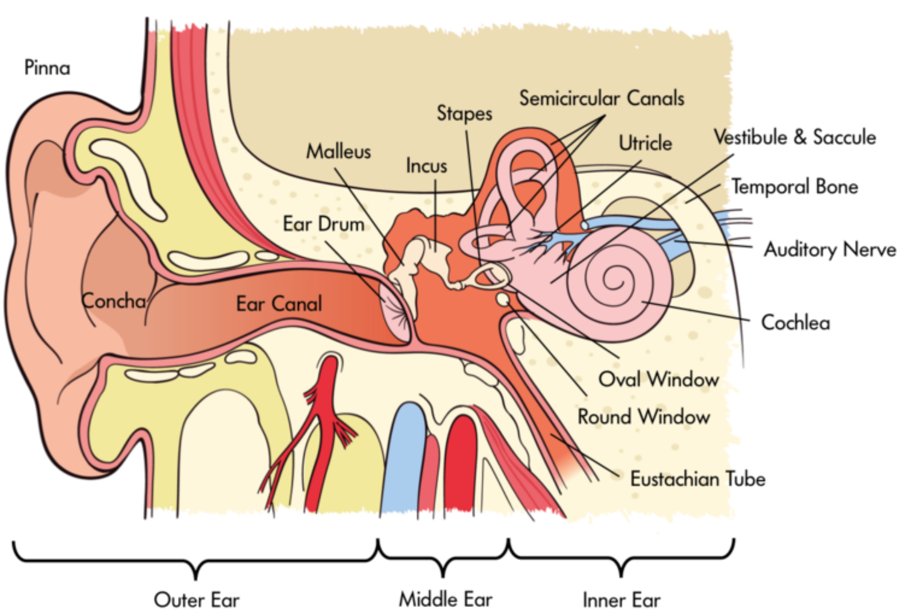


Figure 1: Diagram of the outer, middle, and inner ear.

An Overview of the Auditory System

Our auditory perception depends on how sound is processed through the ear. The ear can be divided into three main parts—the outer, middle, and inner ear (see Figure 1). The outer ear consists of the **pinna** (the visible part of the ear, with all its unique folds and bumps), the ear canal (or auditory meatus), and the **tympanic membrane**. Of course, most of us have two functioning ears, which turn out to be particularly useful when we are trying to figure out where a sound is coming from. As discussed below in the section on spatial hearing, our brain can compare the subtle differences in the signals at the two ears to localize sounds in space. However, this trick does not always help: for instance, a sound directly in front or directly behind you will not produce a difference between the ears. In these cases, the filtering produced by the pinnae helps us localize sounds and resolve potential front-back and up-down confusions. More generally, the folds and bumps of the pinna produce distinct peaks and dips in the frequency response that depend on the location of the sound source. The brain then learns to associate certain patterns of spectral peaks and dips with certain spatial locations. Interestingly, this learned association remains malleable, or plastic, even in adulthood. For instance, a study that altered the pinnae using molds found that people could learn to use their "new" ears accurately within a matter of a few weeks (Hofman, Van Riswick, & Van Opstal, 1998). Because of the small size of the pinna, these kinds of acoustic cues are only found at high frequencies, above about 2 kHz. At lower frequencies, the sound is basically unchanged whether it comes from above, in front, or below. The ear canal itself is a tube that helps to amplify sound in the region from about 1 to 4 kHz—a region particularly important for speech communication.

The middle ear consists of an air-filled cavity, which contains the middle-ear bones, known as the incus, malleus, and stapes, or anvil, hammer, and stirrup, because of their respective shapes. They have the distinction of being the smallest bones in the body. Their primary function is to transmit the vibrations from the tympanic membrane to the oval window of the cochlea and, via a form of lever action, to better match the impedance of the air surrounding the tympanic membrane with that of the fluid within the cochlea.

The inner ear includes the cochlea, encased in the temporal bone of the skull, in which the mechanical vibrations of sound are transduced into neural signals that are processed by the brain. The **cochlea** is a spiral-shaped structure that is filled with fluid. Along the length of the spiral runs the basilar membrane, which vibrates in response to the pressure differences produced by vibrations of the oval window. Sitting on the basilar membrane is the organ of Corti, which runs the entire length of the basilar membrane from the base (by the oval window) to the apex (the "tip" of the spiral). The organ of Corti includes three rows of outer hair cells and one row of inner hair cells. The hair cells sense the vibrations by way of their tiny hairs, or stereocillia. The outer hair cells seem to function to mechanically amplify the sound-induced vibrations, whereas the inner hair cells form synapses with the auditory nerve and transduce those vibrations into action potentials, or neural spikes, which are transmitted along the auditory nerve to higher centers of the auditory pathways.

One of the most important principles of hearing—frequency analysis—is established in the cochlea. In a way, the action of the cochlea can be likened to that of a prism: the many frequencies that make up a complex sound are broken down into their constituent frequencies, with low frequencies creating maximal basilar-membrane vibrations near the apex of the cochlea and high frequencies creating maximal basilar-membrane vibrations nearer the base of the cochlea. This decomposition of sound into its constituent frequencies, and the frequency-to-place mapping, or "tonotopic" representation, is a major organizational principle of the auditory system, and is maintained in the neural representation of sounds all the way from the cochlea to the primary auditory cortex. The decomposition of sound into its constituent frequency components is part of what allows us to hear more than one sound at a time. In addition to representing frequency by place of excitation within the cochlea, frequencies are also represented by the timing of spikes within the auditory nerve. This property, known as "phase locking," is crucial in comparing time-of-arrival differences of waveforms between the two ears (see the section on spatial hearing, below).

Unlike vision, where the primary visual cortex (or V1) is considered an early stage of processing, auditory signals go through many stages of processing before they reach the primary auditory cortex, located in the temporal lobe. Although we have a fairly good understanding of the electromechanical properties of the cochlea and its various structures, our understanding of the processing accomplished by higher stages of the auditory pathways remains somewhat sketchy. With the possible exception of spatial localization and neurons tuned to certain locations in space (Harper & McAlpine, 2004; Knudsen & Konishi, 1978), there is very little consensus on the how, what, and where of auditory feature extraction and representation. There is evidence for a "pitch center" in the auditory cortex from both human neuroimaging studies (e.g., Griffiths, Buchel, Frackowiak, & Patterson, 1998; Penagos, Melcher, & Oxenham, 2004) and single-unit physiology studies (Bendor & Wang, 2005), but even here there remain some questions regarding whether a single area of cortex is responsible for coding single features, such as pitch, or whether the code is more distributed (Walker, Bizley, King, & Schnupp, 2011).

Audibility, Masking, and Frequency Selectivity

Overall, the human cochlea provides us with hearing over a very wide range of frequencies. Young people with normal hearing are able to perceive sounds with frequencies ranging from about 20 Hz all the way up to 20 kHz. The range of intensities we can perceive is also impressive: the quietest sounds we can hear in the medium-frequency range (between about 1 and 4 kHz) have a sound intensity that is about a factor of 1,000,000,000,000 less intense than the loudest sound we can listen to without incurring rapid and permanent hearing loss. In part because of this enormous dynamic range, we tend to use a logarithmic scale, known as decibels (dB), to describe sound pressure or intensity. On this scale, 0 dB sound pressure level (SPL) is defined as 20 micro-Pascals (μPa), which corresponds roughly to the quietest perceptible sound level, and 120 dB SPL is considered dangerously loud.

When the frequency content of different sounds overlaps, masking occurs. Less intense sounds become difficult or impossible to hear because more intense sounds dominate and interfere. Crowded restaurants or busy city streets full of traffic are typical examples of places where certain sounds can "swamp" others. [Image: Peter van der Sluijs, https://goo.gl/K8L4c0, CC BY-SA 3.0, https://goo.gl/eLCn2O]

Masking is the process by which the presence of one sound makes another sound more difficult to hear. We all encounter masking in our everyday lives, when we fail to hear the phone ring while we are taking a shower, or when we struggle to follow a conversation in a noisy restaurant. In general, a more intense sound will mask a less intense sound, provided certain conditions are met. The most important condition is that the frequency content of the sounds overlap, such that the activity in the cochlea produced by a masking sound “swamps” that produced by the target sound. Another type of masking, known as “suppression,” occurs when the response to the masker reduces the neural (and in some cases, the mechanical) response to the target sound. Because of the way that filtering in the cochlea functions, low-frequency sounds are more likely to mask high frequencies than vice versa, particularly at high sound intensities. This asymmetric aspect of masking is known as the “upward spread of masking.” The loss of sharp cochlear tuning that often accompanies cochlear damage leads to broader filtering and more masking—a physiological phenomenon that is likely to contribute to the difficulties experienced by people with hearing loss in noisy environments (Moore, 2007).

Although much masking can be explained in terms of interactions within the cochlea, there are other forms that cannot be accounted for so easily, and that can occur even when interactions within the cochlea are unlikely. These more central forms of masking come in different forms, but have often been categorized together under the term "informational masking" (Durlach et al., 2003; Watson & Kelly, 1978). Relatively little is known about the causes of informational masking, although most forms can be ascribed to a perceptual "fusion" of the masker and target sounds, or at least a failure to segregate the target from the masking sounds. Also relatively little is known about the physiological locus of informational masking, except that at least some forms seem to originate in the auditory cortex and not before (Gutschalk, Micheyl, & Oxenham, 2008).

Spatial Hearing

In contrast to vision, we have a 360° field of hearing. Our auditory acuity is, however, at least an order of magnitude poorer than vision in locating an object in space. Consequently, our auditory localization abilities are most useful in alerting us and allowing us to orient towards sources, with our visual sense generally providing the finer-grained analysis. Of course, there are differences between species, and some, such as barn owls and echolocating bats, have developed highly specialized sound localization systems.

Humans are able to locate sound in space to determine whether the source is in front of us or behind us, or whether it is elevated or below us. [Photo: David Goehring, https://goo.gl/UOLZpB, CC BY 2.0, https://goo.gl/BRvSA7]

Our ability to locate sound sources in space is an impressive feat of neural computation. The two main sources of information both come from a comparison of the sounds at the two ears. The first is based on **interaural time differences (ITD)** and relies on the fact that a sound source on the left will generate sound that will reach the left ear slightly before it reaches the right ear. Although sound is much slower than light, its speed still means that the time of arrival differences between the two ears is a fraction of a millisecond. The largest ITD we encounter in the real world (when sounds are directly to the left or right of us) are only a little over half a millisecond. With some practice, humans can learn to detect an ITD of between 10 and 20 μs (i.e., 20 *millionths* of a second) (Klump & Eady, 1956).

The second source of information is based in **interaural level differences (ILDs)**. At higher frequencies (higher than about 1 kHz), the head casts an acoustic "shadow," so that when a sound is presented from the left, the sound level at the left ear is somewhat higher than the sound level at the right ear. At very high frequencies, the ILD can be as much as 20 dB, and we are sensitive to differences as small as 1 dB.

As mentioned briefly in the discussion of the outer ear, information regarding the elevation of a sound source, or whether it comes from in front or behind, is contained in high-frequency spectral details that result from the filtering effects of the pinnae.

In general, we are most sensitive to ITDs at low frequencies (below about 1.5 kHz). At higher frequencies we can still perceive changes in timing based on the slowly varying temporal envelope of the sound but not the temporal fine structure (Bernstein & Trahiotis, 2002; Smith, Delgutte, & Oxenham, 2002), perhaps because of a loss of neural phase-locking to the temporal fine structure at high frequencies. In contrast, ILDs are most useful at high frequencies, where the head shadow is greatest. This use of different acoustic cues in different frequency regions led to the classic and very early "duplex theory" of sound localization (Rayleigh, 1907). For everyday sounds with a broad frequency spectrum, it seems that our perception of spatial location is dominated by interaural time differences in the low-frequency temporal fine structure (Macpherson & Middlebrooks, 2002).

As with vision, our perception of distance depends to a large degree on context. If we hear someone shouting at a very low sound level, we infer that the shouter must be far away, based on our knowledge of the sound properties of shouting. In rooms and other enclosed locations, the reverberation can also provide information about distance: As a speaker moves further away, the direct sound level decreases but the sound level of the reverberation remains about the same; therefore, the ratio of direct-to-reverberant energy decreases (Zahorik & Wightman, 2001).

Auditory Scene Analysis

There is usually more than one sound source in the environment at any one time—imagine talking with a friend at a café, with some background music playing, the rattling of coffee mugs behind the counter, traffic outside, and a conversation going on at the table next to yours. All these sources produce sound waves that combine to form a single complex waveform at the eardrum, the shape of which may bear very little relationship to any of the waves produced by the individual sound sources. Somehow the auditory system is able to break down, or decompose, these complex waveforms and allow us to make sense of our acoustic environment by forming separate auditory "objects" or "streams," which we can follow as the sounds unfold over time (Bregman, 1990).

A number of heuristic principles have been formulated to describe how sound elements are grouped to form a single object or segregated to form multiple objects. Many of these originate from the early ideas proposed in vision by the so-called Gestalt psychologists, such as Max Wertheimer. According to these rules of thumb, sounds that are in close proximity, in time or frequency, tend to be grouped together. Also, sounds that begin and end at the same time tend to form a single auditory object. Interestingly, spatial location is not always a strong or reliable grouping cue, perhaps because the location information from individual frequency components is often ambiguous due to the effects of reverberation. Several studies have looked into the relative importance of different cues by "trading off" one cue against another. In some cases, this has led to the discovery of interesting auditory illusions, where melodies that are not present in the sounds presented to either ear emerge in the perception (Deutsch, 1979), or where a sound element is perceptually "lost" in competing perceptual organizations (Shinn-Cunningham, Lee, & Oxenham, 2007).

More recent attempts have used computational and neutrally based approaches to uncover the mechanisms of auditory scene analysis (e.g., Elhilali, Ma, Micheyl, Oxenham, & Shamma, 2009), and the field of computational auditory scene analysis (CASA) has emerged in part as an effort to move towards more principled, and less heuristic, approaches to understanding the parsing and perception of complex auditory scenes (e.g., Wang & Brown, 2006). Solving this problem will not only provide us with a better understanding of human auditory perception, but may provide new approaches to "smart" hearing aids and cochlear implants, as well as automatic speech recognition systems that are more robust to background noise.

An infant with a cochlear implant. [Image: Bjorn Knetsch, https://goo.gl/J2wCvJ, CC BY 2.0, https://goo.gl/BRvSA7]

Conclusion

Hearing provides us with our most important connection to the people around us. The intricate physiology of the auditory system transforms the tiny variations in air pressure that reach our ear into the vast array of auditory experiences that we perceive as speech, music, and sounds from the environment around us. We are only beginning to understand the basic principles of neural coding in higher stages of the auditory system, and how they relate to perception. However, even our rudimentary understanding has improved the lives of hundreds of thousands through devices such as cochlear implants, which re-create some of the ear's functions for people with profound hearing loss.

Outside Resources

Audio: Auditory Demonstrations from Richard Warren's lab at the University of Wisconsin, Milwaukee
http://www4.uwm.edu/APL/demonstrations.html

Audio: Auditory Demonstrations. CD published by the Acoustical Society of America (ASA). You can listen to the demonstrations here
http://www.feilding.net/sfuad/musi3012-01/demos/audio/

Web: Demonstrations and illustrations of cochlear mechanics can be found here
http://lab.rockefeller.edu/hudspeth/graphicalSimulations

Web: More demonstrations and illustrations of cochlear mechanics
http://www.neurophys.wisc.edu/animations/

Discussion Questions

1. Based on the available acoustic cues, how good do you think we are at judging whether a low-frequency sound is coming from in front of us or behind us? How might we solve this problem in the real world?
2. Outer hair cells contribute not only to amplification but also to the frequency tuning in the cochlea. What are some of the difficulties that might arise for people with cochlear hearing loss, due to these two factors? Why do hearing aids not solve all these problems?
3. Why do you think the auditory system has so many stages of processing before the signals reach the auditory cortex, compared to the visual system? Is there a difference in the speed of processing required?

Vocabulary

Cochlea
Snail-shell-shaped organ that transduces mechanical vibrations into neural signals.

Interaural differences
Differences (usually in time or intensity) between the two ears.

Pinna
Visible part of the outer ear.

Tympanic membrane
Ear drum, which separates the outer ear from the middle ear.

References

Attneave, F., & Olson, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 84, 147–166.

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436, 1161–1165.

Bernstein, L. R., & Trahiotis, C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli." *Journal of the Acoustical Society of America*, 112, 1026–1036.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Deutsch, D. (1979). Binaural integration of melodic patterns. *Perception & Psychophysics*, 25, 399–405.

Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, 113, 2984–2987.

Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. (2009). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron*, 61, 317–329.

Fechner, G. T. (1860). *Elemente der Psychophysik* (Vol. 1). Leipzig, Germany: Breitkopf und Haertl.

Griffiths, T. D., Buchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, 1, 422–427.

Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, 6, 1156–1165 (e1138).

Harper, N. S., & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature*, 430, 682–686.

Hofman, P. M., Van Riswick, J. G. A., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience*, 1, 417–421.

ISO. (2003). *ISO:226 Acoustics - Normal equal-loudness-level contours*. Geneva, Switzerland: International Organization for Standardization.

Klump, R. G., & Eady, H. R. (1956). Some measurements of interaural time difference thresholds. *Journal of the Acoustical Society of America*, 28, 859–860.

Knudsen, E. I., & Konishi, M. (1978). A neural map of auditory space in the owl. *Science*, 200, 795–797.

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *Journal of the Acoustical Society of America*, 111, 2219–2236.

Moore, B. C. J. (2007). *Cochlear hearing loss: Physiological, psychological, and technical issues*. Chichester: Wiley.

Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45, 224–240.

Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., & Santurette, S. (2011). Pitch perception beyond the traditional existence region of pitch. *Proceedings of the National Academy of Sciences USA*, 108, 7629–7634.

Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in non-primary human auditory cortex revealed with fMRI. *Journal of Neuroscience*, 24, 6810–6815.

Pressnitzer, D., Patterson, R. D., & Krumbholz, K. (2001). The lower limit of melodic pitch. *Journal of the Acoustical Society of America*, 109, 2074–2084.

Rayleigh, L. (1907). On our perception of sound direction. *Philosophical Magazine*, 13, 214–232.

Risset, J. C., & Wessel, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 113–168): Academic Press.

Shinn-Cunningham, B. G., Lee, A. K., & Oxenham, A. J. (2007). A sound element gets lost in perceptual competition. *Proceedings of the National Academy of Sciences USA*, 104, 12223–12227.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87–90.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181.

Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. W. (2011). Multiplexed and robust representations of sound features in auditory cortex. *Journal of Neuroscience*, 31, 14565–14576.

Wang, D., & Brown, G. J. (Eds.). (2006). *Computational auditory scene analysis: Principles, algorithms, and applications*. Hoboken, NJ: Wiley.

Watson, C. S., & Kelly, W. J. (1978). Informational masking in auditory patterns. *Journal of the Acoustical Society of America*, 64, S39.

Zahorik, P., & Wightman, F. L. (2001). Loudness constancy with varying sound source distance. *Nature Neuroscience*, 4, 78–83.

de Cheveigné, A. (2005). Pitch perception models. In C. J. Plack, A. J. Oxenham, A. N. Popper,

& R. Fay (Eds.), *Pitch: Neural coding and perception* (pp. 169–233). New York, NY: Springer Verlag.

14

Taste and Smell

Linda Bartoshuk & Derek Snyder

Humans are omnivores (able to survive on many different foods). The omnivore's dilemma is to identify foods that are healthy and avoid poisons. Taste and smell cooperate to solve this dilemma. Stimuli for both taste and smell are chemicals. Smell results from a biological system that essentially permits the brain to store rough sketches of the chemical structures of odor stimuli in the environment. Thus, people in very different parts of the world can learn to like odors (paired with calories) or dislike odors (paired with nausea) that they encounter in their worlds. Taste information is preselected (by the nature of the receptors) to be relevant to nutrition. No learning is required; we are born loving sweet and hating bitter. Taste inhibits a variety of other systems in the brain. Taste damage releases that inhibition, thus intensifying sensations like those evoked by fats in foods. Ear infections and tonsillectomies both can damage taste. Adults who have experienced these conditions experience intensified sensations from fats and enhanced palatability of high-fat foods. This may explain why individuals who have had ear infections or tonsillectomies tend to gain weight.

Learning Objectives

- Explain the salient properties of taste and smell that help solve the omnivore's dilemma.
- Distinguish between the way pleasure/displeasure is produced by smells and tastes.
- Explain how taste damage can have extensive unexpected consequences.

The Omnivore's Dilemma

Humans are **omnivores**. We can survive on a wide range of foods, unlike species, such as koalas, that have a highly specialized diet (for koalas, eucalyptus leaves). With our amazing dietary range comes a problem: the omnivore's dilemma (Pollan, 2006; Rozin & Rozin, 1981). To survive, we must identify healthy food and avoid poisons. The senses of taste and smell cooperate to give us this ability. Smell also has other important functions in lower animals (e.g., avoid predators, identify sexual partners), but these functions are less important in humans. This module will focus on the way taste and smell interact in humans to solve the omnivore's dilemma.

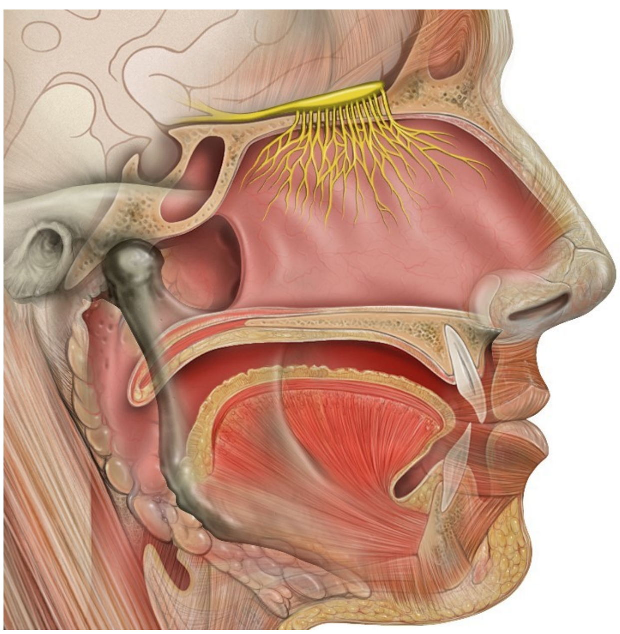
By just examining our teeth, it is apparent that we are animals who evolved to eat a wide range of foods. [Image: Justin Mclean, https://goo.gl/ffk7ZV, CC BY-NC 2.0, https://goo.gl/VnKlK8]

Taste and Smell Anatomy

Taste (**gustation**) and smell (**olfaction**) are both chemical senses; that is, the stimuli for these senses are chemicals. The more complex sense is olfaction. Olfactory receptors are complex proteins called G protein-coupled receptors (GPCRs). These structures are proteins that weave back and forth across the membranes of olfactory cells seven times, forming structures outside the cell that sense odorant molecules and structures inside the cell that activate the neural message ultimately conveyed to the brain by olfactory neurons. The structures that sense odorants can be thought of as tiny binding pockets with sites that respond to active parts of molecules (e.g., carbon chains). There are about 350 functional olfactory genes in humans; each gene expresses a particular kind of olfactory receptor. All olfactory receptors of a given kind project to structures called glomeruli (paired clusters of cells found on both sides of the brain). For a single molecule, the pattern of activation across the glomeruli paints a picture of the chemical structure of the molecule. Thus, the olfactory system can identify a vast array of chemicals present in the environment. Most of the odors we encounter are actually mixtures of chemicals (e.g., bacon odor). The olfactory system creates an image for the mixture and stores it in memory just as it does for the odor of a single molecule (Shepherd, 2005).

Taste is simpler than olfaction. Bitter and sweet utilize GPCRs, just as olfaction does, but the number of different receptors is much smaller. For bitter, 25 receptors are tuned to different chemical structures (Meyerhof et al., 2010). Such a system allows us to sense many different poisons.

Sweet is even simpler. The primary sweet receptor is composed of two different G protein-coupled receptors; each of these two proteins ends in large structures reminiscent of Venus flytraps. This complex receptor has multiple sites that can bind different structures. The Venus flytrap endings open so that even some very large molecules can fit inside and stimulate the receptor.



Although smell plays a less integral role in our lives than it does for other animals, it is highly important in determining taste. In fact, if you plug your nose while eating chocolate, you will have a difficult time distinguishing it from any other kind of sweet. [Image: Patrick J. Lynch, https://goo.gl/d5CIff, CC BY 2.5, https://goo.gl/0QtWcf]

Bitter is inclusive (i.e., multiple receptors tuned to very different chemical structures feed into common neurons). Sweet is exclusive. There are many sugars with similar structures, but only three of these are particularly important to humans (sucrose, glucose, and fructose). Thus, our sweet receptor tunes out most sugars, leaving only the most important to stimulate the sweet receptor. However, the ability of the sweet receptor to respond to some non-sugars presents us with one of the great mysteries of taste. Several non-sugar molecules can stimulate the primary sweet receptor (e.g., saccharine, aspartame, cyclamate). These have given rise to the artificial sweetener industry, but their biological significance is unknown. What biological purpose is served by allowing these non-sugar molecules to stimulate the primary sweet receptor?

Some would have us believe that artificial sweeteners are a boon to those who want to lose weight. It seems like a no-brainer. Sugars have calories; saccharin does not. Theoretically, if we replace sugar with saccharin in our diets, we will lose weight. In fact, recent work showed that rats actually gained weight when saccharin was substituted for glucose (Swithers & Davidson, 2008). It turns out that substituting saccharin for sugar can increase appetite so more is eaten later. In addition, eating artificial sweeteners appears to alter metabolism, thus making losing weight even harder. So why did nature give us artificial sweeteners? We don't know.

One more mystery about sweet deserves comment. The discovery of the sweet receptor was met with great excitement because many investigators had searched for it for years. The fact that this complex receptor had multiple sites to which different molecules could bind explained why many different molecules taste sweet. However, this is actually a serious problem. No matter what molecule stimulates this receptor, the neural output from that receptor is the same. This would mean that the sweetness of all sweet substances would have to be the same. Yet artificial sweeteners do not taste exactly like sugar. The answer may lie in the fact that one of the two proteins that makes up the receptor can act alone, but only strong concentrations of sugar stimulate this isolated protein receptor. This permits the brain to distinguish between the sweetness of sugar and the sweetness of non-sugar molecules.

Salty and sour are the simplest tastes; these stimuli ionize (break into positively and negatively charged particles). The first event in the transduction series is the movement of the positively charged particle through channels in the taste cell membrane (Chaudhari & Roper, 2010).

Just how biologically ingrained is our love for sweets? Newborn infants immediately love sweet tastes, while the taste for salty foods takes longer to develop. [Image: shingleback, https://goo.gl/fbUH3r, CC BY 2.0, https://goo.gl/BRvSA7]

Solving the omnivore's dilemma: Taste affect is hard-wired

The pleasure associated with sweet and salty and the displeasure associated with sour and bitter are hard-wired in the brain. Newborns love sweet (taste of mother's milk) and hate bitter (poisons) immediately. The receptors mediating salty taste are not mature at birth in humans, but when they are mature a few weeks after birth, the baby likes dilute salt (although more concentrated salt will evoke stinging sensations that will be avoided). Sour is generally disliked (protecting against tissue damage from acid?), but to the amazement of many parents, some young children appear to actually like the sour candies available today; this may be related to the breadth of their experience with fruits (Liem & Mennella, 2003). This hard-wired affect is the most salient characteristic of taste and this is why we classify only those taste qualities with hard-wired affect as "basic tastes."

Another contribution to the omnivore's dilemma: Olfactory affect is learned

The biological functions of olfaction depend on how odors enter our noses. Sniffing brings odorants through our nostrils. The odorants hit the turbinate bones and a puff of the odorized air rises to the top of the nasal cavity, where it goes through a narrow opening (the olfactory cleft) and arrives at the olfactory mucosa (the tissue that houses the olfactory receptors). Technically, this is called "**orthonasal olfaction**." Orthonasal olfaction tells us about the world external to our bodies.

When we chew and swallow food, the odorants emitted by the food are forced up behind the palate (roof of the mouth) and enter our noses from the back; this is called "**retronasal olfaction**." Ortho and retronasal olfaction involve the same odor molecules and the same olfactory receptors; however, the brain can tell the difference between the two and does not send the input to the same areas. Retronasal olfaction and taste project to some common areas where they are presumably integrated into flavor. Flavors tell us about the food we are eating.

If retronasal olfaction is paired with nausea, the food evoking the retronasal olfactory sensation becomes disliked. If retronasal olfaction is paired with situations the brain deems valuable (calories, sweet taste, pleasure from other sources, etc.), the food evoking that sensation becomes liked. These are called **conditioned aversions and preferences** (Rozin & Vollmecke, 1986).

Those who have experienced a conditioned aversion may have found that the dislike (even disgust) evoked when a flavor is paired with nausea can generalize to the smell of the food alone (orthonasal olfaction). Some years ago, Jeremy Wolfe and Linda Bartoshuk surveyed conditioned aversions among college students and staff that had resulted from consuming foods/beverages associated with nausea (Bartoshuk & Wolfe, 1990). In 29% of the aversions, subjects reported that even the smell of the food/beverage had become aversive. Other properties of food objects can become aversive as well. In one unusual case, an aversion to

There are few associations more powerful than the taste of food and an experience of sickness. In fact, many people will go their whole lives without ever trying a food again that once made them sick. [Image: James Palinsad, https://goo.gl/r2Bph4, CC BY-ND 2.0, https://goo.gl/VnHMcZ]

cheese crackers generalized to vanilla wafers apparently because the containers were similar. Conditioned aversions function to protect us from ingesting a food that our brains associate with illness. Conditioned preferences are harder to form, but they help us learn what is safe to eat.

Is the affect associated with olfaction ever hard-wired? Pheromones are said to be olfactory molecules that evoke specific behaviors. Googling "human pheromone" will take you to websites selling various sprays that are supposed to make one more sexually appealing. However, careful research does not support such claims in humans or any other mammals (Doty, 2010). For example, amniotic fluid was at one time believed to contain a pheromone that attracted rat pups to their mother's nipples so they could suckle. Early interest in identifying the molecule that acted as that pheromone gave way to understanding that the behavior was learned when a novel odorant, citral (which smells like lemons), was easily substituted for amniotic fluid (Pedersen, Williams, & Blass, 1982).

Central interactions: Key to understanding taste damage

The integration of retronasal olfaction and taste into flavor is not the only central interaction among the sensations evoked by foods. These integrations in most cases serve important biological functions, but occasionally they go awry and lead to clinical pathologies.

Taste is mediated by three cranial nerves; these are bilateral nerves, each of which innervates one side of the mouth. Since they do not connect in the peripheral nervous system, interactions across the midline must occur in the brain. Incidentally, studying interactions across the midline is a classic way to draw inferences about central interactions. Insights from studies of this type were very important to understanding central processes long before we had direct imaging of brain function.

Taste on the anterior two thirds of the tongue (the part you can stick out) is mediated by the chorda tympani nerve; taste on the posterior one third (the part that stays attached) is mediated by the glossopharyngeal nerve. Taste buds are tiny clusters of cells (like the segments of an orange) that are buried in the tissue of some papillae, the structures that give the tongue its bumpy appearance. Filiform papillae are the smallest and are distributed all over the tongue; they have no taste buds. In species like the cat, the filiform papillae are shaped like small spoons and help the cat hold liquids on the tongue while lapping (try lapping from a dish and you will see how hard it is without those special filiform papillae). Fungiform papillae (given this name because they resemble small button mushrooms) are larger circular structures on the anterior tongue (innervated by the chorda tympani). They contain about six taste buds each. Fungiform papillae can be seen with the naked eye, but swabbing blue food coloring on the tongue helps. The fungiform papillae do not stain as well as the rest of the tongue so they look like pink circles against a blue background. On some tongues, the spacing of fungiform papillae is like polka dots. Other tongues can have 10 times as many fungiform papillae, spaced so closely that there is little space between them. There is a connection between the density of fungiform papillae and the perception of taste. Those who experience the most intense taste sensations (we call them supertasters) tend to have the most fungiform papillae. Incidentally, this is a rare example in sensory processes of visible anatomical variation that correlates with function. We can look at the tongues of a variety of individuals and predict which of them will experience the most intense taste sensations.

Approximately 25% of people are "non-tasters," 50% are "medium tasters," and another 25% are "super tasters." The better the "taster" one is, the less he or she can typically tolerate spicy foods, salt, and the taste of alcohol. [Image: Rosie Rogers, https://goo.gl/j7SepS, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

The structures that house taste buds innervated by the glossopharyngeal nerve are called circumvallate papillae. They are relatively large structures arrayed in an inverted V shape across the back of the tongue. Each of them looks like a small island surrounded by a moat.

Taste nerves project to the brain, where they send inhibitory signals to one another. One of the biological consequences of this inhibition is taste constancy. Damage to one nerve reduces taste input but also reduces inhibition on the other nerves (Bartoshuk et al 2005). That release of inhibition intensifies the central neural signals from the undamaged nerves, thereby maintaining whole mouth function. Interestingly, this release of inhibition can be so powerful that it actually increases whole mouth taste. The small effect of limited taste damage is one of the earliest clinical observations. In 1825, Brillat-Savarin described in his book *The Physiology of Taste* an interview with an ex-prisoner who had suffered a horrible punishment: amputation of his tongue. "This man, whom I met in Amsterdam, where he made his living by running errands, had had some education, and it was easy to communicate with him by writing. After I had observed that the forepart of his tongue has been cut off clear to the ligament, I asked him if he still found any flavor in what he ate, and if his sense of taste had survived the cruelty to which he had been subjected. He replied that … he still possessed the ability to taste fairly well" (Brillat-Savarin, 1971, pg. 35). This injury damaged the chorda tympani but spared the glossopharyngeal nerve.

We now know that taste nerves not only inhibit one another but also inhibit other oral sensations. Thus, taste damage can intensify oral touch (fats) and oral burn (chilis). In fact, taste damage appears to be linked to pain in general. Consider an animal injured in the wild. If pain reduced eating, its chance of survival would be diminished. However, nature appears to have wired the brain such that taste input inhibits pain. Eating is reinforced and the animal's chances of survival increase.

Taste damage and weight gain

The effects of taste damage depend on the extent of damage. If only one taste nerve is damaged, then release of inhibition occurs. If the damage is extensive enough, function is lost with one possible exception. Preliminary data suggest that the more extensive the damage to taste, the greater the intensification of pain; this is obviously of clinical interest.

Damage to a single taste nerve can intensify oral touch (e.g., the creamy, viscous sensations evoked by fats). Perhaps most surprising, damage to a single taste nerve can intensify retronasal olfaction; this may occur as a secondary result from the intensification of whole mouth taste.

These sensory changes can alter the palatability of foods; in particular, high-fat foods can be rendered more palatable. Thus one of the first areas we examined was the possibility that mild taste damage could lead to increases in body mass index. Middle ear infections (otitis media) can damage the chorda tympani nerve; a tonsillectomy can damage the glossopharyngeal nerve. Head trauma damages both nerves, although it tends to take its greatest toll on the chorda tympani nerve. All of these clinical conditions increase body mass index in some individuals. More work is needed, but we suspect a link between the intensification of fat sensations, enhancement of palatability of high-fat foods, and weight gain.

Outside Resources

Video: Inside the Psychologists Studio with Linda Bartoshuk
http://www.youtube.com/watch?v=fO76jgepo74

Video: Linda Bartoshuk at Nobel Conference 46
http://www.youtube.com/watch?v=I4tbg_yb4Ms

Video: Test your tongue: the science of taste
http://www.youtube.com/watch?v=KvP7xGBzZNs

Discussion Questions

1. In this module, we have defined "basic tastes" in terms of whether or not a sensation produces hard-wired affect. Can you think of any other definitions of basic tastes?
2. Do you think omnivores, herbivores, or carnivores have a better chance at survival?
3. Olfaction is mediated by one cranial nerve. Taste is mediated by three cranial nerves. Why do you think evolution gave more nerves to taste than to smell? What are the consequences of this?

Vocabulary

Conditioned aversions and preferences
Likes and dislikes developed through associations with pleasurable or unpleasurable sensations.

Gustation
The action of tasting; the ability to taste.

Olfaction
The sense of smell; the action of smelling; the ability to smell.

Omnivore
A person or animal that is able to survive by eating a wide range of foods from plant or animal origin.

Orthonasal olfaction
Perceiving scents/smells introduced via the nostrils.

Retronasal olfaction
Perceiving scents/smells introduced via the mouth/palate.

References

Bartoshuk, L. M., & Wolfe, J. M. (1990). Conditioned taste aversions in humans: Are they olfactory aversions? (abstract). *Chemical Senses, 15*, 551.

Bartoshuk, L. M., Snyder, D. J., Grushka, M., Berger, A. M., Duffy, V. B,, & Kveton, J. F. (2005). Taste damage: previously unsuspected consequences. *Chemical Senses, 30*(Suppl. 1), i218–i219

Brillat-Savarin, J. A. (1825). *The physiology of taste* (M.F.K. Fisher, Trans., 1971). New York, NY: Alfred A. Knopf.

Chaudhari, N., & Roper, S. D. (2010). The cell biology of taste. *Journal of Cell Biology, 190*, 285–296.

Doty, R. L. (2010). *The great pheromone myth*. Baltimore, MD: Johns Hopkins University Press.

Liem, D. G., & Mennella, J. A. (2003). Heightened sour preferences during childhood. *Chemical Senses, 28*(2), 173–180.

Meyerhof, W., Batram, C., Kuhn, C., Brockhoff, A., Chudoba, E., Bufe, B., . . . Behrens, M. (2010). The molecular receptive ranges of human TAS2R bitter taste receptors. *Chemical Senses, 35*, 157–170.

Pedersen, P.E., Williams, C.L., & Blass, E.M. (1982). Activation and odor conditioning of suckling behavior in 3-day-old albino rats. *Journal of Experimental Psychology: Animal Behavior Processes, 8*(4), 329–341.

Pollan, M. (2006). *The omnivore's dilemma*. New York, NY: Penguin Books.

Rozin, E., & Rozin, P. (1981). Culinary themes and variations. *Natural History, 90*, 6–14.

Rozin, P., & Vollmecke, T.A. (1986). Food likes and dislikes. *Annual Review of Nutrition, 6*, 433-456.

Shepherd, G. M. (2005). Outline of a theory of olfactory processing and its relevance to humans. *Chemical Senses, 30*(Suppl 1), i3-i5.

Swithers, S. E., & Davidson, T. L. (2008). A role for sweet taste: Calorie predictive relations in energy regulation by rats. *Behavioral Neuroscience, 122*(1), 161-173.

15
Touch and Pain

Guro E. Løseth, Dan-Mikael Ellingson & Siri Leknes

The sensory systems of touch and pain provide us with information about our environment and our bodies that is often crucial for survival and well-being. Moreover, touch is a source of pleasure. In this module, we review how information about our environment and our bodies is coded in the periphery and interpreted by the brain as touch and pain sensations. We discuss how these experiences are often dramatically shaped by top-down factors like motivation, expectation, mood, fear, stress, and context. When well-functioning, these circuits promote survival and prepare us to make adaptive decisions. Pathological loss of touch can result in perceived disconnection from the body, and insensitivity to pain can be very dangerous, leading to maladaptive hazardous behavior. On the other hand, chronic pain conditions, in which these systems start signaling pain in response to innocuous touch or even in the absence of any observable sensory stimuli, have tremendous negative impact on the lives of the affected. Understanding how our sensory-processing mechanisms can be modulated psychologically and physiologically promises to help researchers and clinicians find new ways to alleviate the suffering of chronic-pain patients

Learning Objectives

- Describe the transduction of somatosensory signals: The properties of the receptor types as well as the difference in the properties of C-afferents and A-afferents and what functions these are thought to have.
- Describe the social touch hypothesis and the role of affective touch in development and bonding.
- Explain the motivation–decision model and descending modulation of pain, and give examples on how this circuitry can promote survival.
- Explain how expectations and context affect pain and touch experiences.

- Describe the concept of chronic pain and why treatment is so difficult.

Introduction

Imagine a life free of pain. How would it be—calm, fearless, serene? Would you feel invulnerable, invincible? Getting rid of pain is a popular quest—a quick search for "pain-free life" on Google returns well over 4 million hits—including links to various bestselling self-help guides promising a pain-free life in only 7 steps, 6 weeks, or 3 minutes. Pain management is a billion-dollar market, and involves much more than just pharmaceuticals. Surely a life with no pain would be a better one?

Figure 1A: Patient with HSAN-V genetic mutation affecting pain nerve growth. Severely affected 12-year-old boy with damages to his left knee and ankles. (Minde et al., 2004) [Used with permission]

Well, consider one of the "lucky few": 12-year-old "Thomas" has never felt deep pain. Not even when a fracture made him walk around with one leg shorter than the other, so that the bones of his healthy leg were slowly crushed to destruction underneath the knee joint (see Figure 1A). For Thomas and other members of a large Swedish family, life without pain is a harsh reality because of a mutated gene that affects the growth of the nerves conducting deep pain. Most of those affected suffer from joint damage and frequent fractures to bones in their feet and hands; some end up in wheelchairs even before they reach puberty (Minde et al., 2004). It turns out pain—generally—serves us well.

Living without a sense of touch sounds less attractive than being free of pain—touch is a source of pleasure and essential to how we feel. Losing the sense of touch has severe implications—something patient G. L. experienced when an antibiotics treatment damaged the type of nerves that signal touch from her skin and the position of her joints and muscles. She reported feeling like she'd lost her physical self from her nose down, making her "disembodied"—like she no longer had any connection to the body attached to her head.

If she didn't look at her arms and legs they could just "wander off" without her knowing—initially she was unable to walk, and even after she relearned this skill she was so dependent on her visual attention that closing her eyes would cause her to land in a hopeless heap on the floor. Only light caresses like those from her children's hands can make her feel she has a body, but even these sensations remain vague and elusive (Olausson et al., 2002; Sacks, 1985).

Sensation

Cutaneous Senses of the Skin Connect the Brain to the Body and the Outside World

Touch and pain are aspects of the somatosensory system, which provides our brain with information about our own body (**interoception**) and properties of the immediate external world (**exteroception**) (Craig, 2002). We have somatosensory receptors located all over the body, from the surface of our skin to the depth of our joints. The information they send to the central nervous system is generally divided into four modalities: **cutaneous senses** (senses of the skin), proprioception (body position), kinesthesis (body movement), and **nociception** (pain, discomfort). We are going to focus on the cutaneous senses, which respond to tactile, thermal, and pruritic (itchy) stimuli, and events that cause tissue damage (and hence pain). In addition, there is growing evidence for a fifth modality specifically channeling *pleasant* touch (McGlone & Reilly, 2010).

Different Receptor Types Are Sensitive to Specific Stimuli

The skin can convey many sensations, such as the biting cold of a wind, the comfortable pressure of a hand holding yours, or the irritating itch from a woolen scarf. The different types of information activate specific receptors that convert the stimulation of the skin to electrical nerve impulses, a process called **transduction**. There are three main groups of receptors in our skin: *mechanoreceptors*, responding to mechanical stimuli, such as stroking, stretching, or vibration of the skin; *thermoreceptors*, responding to cold or hot temperatures; and *chemoreceptors*, responding to certain types of chemicals either applied externally or released within the skin (such as histamine from an inflammation). For an overview of the different receptor types and their properties, see Box 1. The experience of *pain* usually starts with activation of **nociceptors**—receptors that fire specifically to potentially tissue-damaging stimuli. Most of the nociceptors are subtypes of either chemoreceptors or mechanoreceptors. When tissue is damaged or inflamed, certain chemical substances are released from the cells, and these substances activate the chemosensitive nociceptors. Mechanoreceptive nociceptors have a high threshold for activation—they respond to mechanical stimulation that is so intense it might damage the tissue.

Box 1. Categories of low-threshold mechanoreceptors*

Identity of receptor	Size of receptor*	Type of skin where found	Speed of adaptation*	Adequate stimulus*
Merkel's disks	Small, sharp borders	Glabrous*	Slow	Pressure
Meissner's corpusles	Small, sharp borders	Glabrous	Rapid	Indentation
Ruffini corpuscles	Large, diffuse borders	Hairy + glabrous	Slow	Stretching
Pacinian corpuscles	Large, diffuse borders	Hairy + glabrous	Rapid	Vibration

* **Adequate stimulus:** The type of stimulus that the receptor is specialized to receive and respond to.

* **Glabrous Skin:** The hairless skin found on our palms and the soles of our feet. This skin has a higher density of receptors of a more complex range, which reflects the fact that we use these areas of our body to actively explore our surroundings and to discriminate tactile properties of objects we're interacting with.

* **Low-threshold mechanoreceptors:** Mechanoreceptors that respond to stimulus that is so light it doesn't threaten to damage the tissue around it. High-threshold mechanoreceptors respond to stimulation of higher intensity, and are a type of nociceptor.

* **Receptive field:** The space of skin or tissue in which stimulation will elicit a response in the receptor. Smaller receptive fields make the receptor more sensitive to details.

* **Speed adaptation:** Slowly adapting mechanoreceptors continue to fire action potentials during sustained stimulation. Rapidly adapting mechanoreceptors fire action poentials in response to stimulus onset and offset (i.e. to stimuli chages), and help detect stimulus movement on the skin.

Action Potentials in the Receptor Cells Travel as Nerve Impulses with Different Speeds

When you step on a pin, this activates a host of mechanoreceptors, many of which are nociceptors. You may have noticed that the sensation changes over time. First you feel a sharp stab that propels you to remove your foot, and only then you feel a wave of more aching pain. The sharp stab is signaled via fast-conducting A-fibers, which project to the somatosensory cortex. This part of the cortex is somatotopically organized—that is, the sensory signals are represented according to where in the body they stem from (see llustration, Figure 2). The unpleasant ache you feel after the sharp pin stab is a separate, simultaneous signal sent from the nociceptors in your foot via thin C-pain or Aδ-fibers to the insular cortex and other brain regions involved in processing of emotion and interoception (see Figure 3a for a schematic representation of this pathway). The experience of stepping on a pin is, in other words, composed by two separate signals: one discriminatory signal allowing us to localize the touch stimulus and distinguish whether it's a blunt or a sharp stab; and one affective signal that lets us know that stepping on the pin is bad. It is common to divide pain into sensory–discriminatory and affective–motivational aspects (Auvray, Myin, & Spence, 2010). This

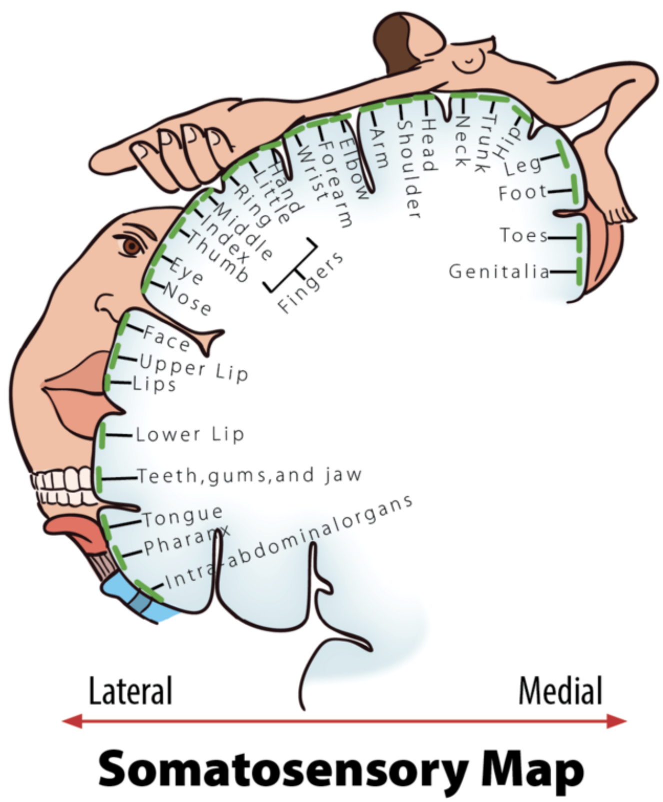


Figure 2: Somatosensory map: Body parts are represented in specific locations on the somatosensory cortex. Representations map out somatotopically, with the feet located medially and shoulders and arms laterally to the interhemispheric fissure. Facial structures are represented in a different location to the scalp and head; the face oriented upside down with the forehead pointing towards the shoulders.

distinction corresponds, at least partly, to how this information travels from the peripheral to the central nervous system and how it is processed in the brain (Price, 2000).

Affective Aspects of Touch Are Important for Development and Relationships

Touch senses are not just there for discrimination or detection of potentially painful events, as Harlow and Suomi (1970) demonstrated in a series of heartbreaking experiments where baby monkeys were taken from their mothers. The infant monkeys could choose between two artificial surrogate mothers—one "warm" mother without food but with a furry, soft cover; and one cold, steel mother with food. The monkey babies spent most of their time clinging to the soft mother, and only briefly moved over to the hard, steel mother to feed, indicating that touch is of "overpowering importance" to the infant (Harlow & Suomi, 1970, p. 161). Gentle touch is central for creating and maintaining social relationships in primates; they groom each other by stroking the fur and removing parasites—an activity important not only for their individual well-being but also for group cohesion (Dunbar, 2010; Keverne, Martensz, & Tuite, 1989). Although people don't groom each other in the same way, gentle touch is important for us, too.

The sense of touch is the first to develop while one is in the womb, and human infants crave touch from the moment they're born. From studies of human orphans, we know that touch is also crucial for human development. In Romanian orphanages where the babies were fed but not given regular attention or physical contact, the children suffered cognitive and neurodevelopmental delay (Simons & Land, 1987). Physical contact helps a crying baby calm down, and the soothing touch a mother gives to her child is thought to reduce the levels of stress hormones such as cortisol. High levels of cortisol have negative effects on neural development, and they can even lead to cell loss (Feldman, Singer, & Zagoory, 2010; Fleming, O'Day, & Kraemer, 1999; Pechtel & Pizzagalli, 2011). Thus, stress reduction through hugs and caresses might be important not only for children's well-being, but also for the development of the infant brain.

The skin senses are similar across species, likely reflecting the evolutionary advantage of being able to tell what is touching you, where it's happening, and whether or not it's likely to cause tissue damage. An intriguing line of touch research suggests that humans, cats, and other animals have a special, evolutionarily preserved system that promotes gentle touch because it carries social and emotional significance.On a peripheral level, this system consists of a subtype of C-fibers that responds not to painful stimuli, but rather to gentle stroking touch —called C-tactile fibers. The firing rate of the C-tactile fibers correlates closely with how pleasant the stroking feels—suggesting they are coding specifically for the gentle caresses typical of social affiliative touch (Löken, Wessberg, Morrison, McGlone, & Olausson, 2009). This finding has led to the **social touch hypothesis**, which proposes that C-tactile fibers form a system for touch perception that supports social bonding (Morrison, Löken, & Olausson, 2010; Olausson, Wessberg, Morrison, McGlone, & Vallbo, 2010). The discovery of the C-tactile system suggests that touch is organized in a similar way to pain; fast-conducting A-fibers contribute to sensory–discriminatory aspects, while thin C-fibers contribute to affective–motivational aspects (Löken, Wessberg, Morrison, McGlone, & Olausson, 2009). However, while these "hard-wired" afferent systems often provide us with accurate information about our environment and our bodies, how we experience touch or pain depends very much on top-down sources like motivation, expectation, mood, fear, and stress.

Modulation

Pain Is Necessary for Survival, but Our Brain Can Stop It if It Needs To

In April 2003, the climber Aron Ralston found himself at the floor of Blue John Canyon in Utah, forced to make an appalling choice: face a slow but certain death—or amputate his right arm. Five days earlier he fell down the canyon—since then he had been stuck with his right arm trapped between an 800-lb boulder and the steep sandstone wall. Weak from lack of food and water and close to giving up, it occurred to him like an epiphany that if he broke the two bones in his forearm he could manage to cut off the rest with his pocket knife. The thought of freeing himself and surviving made him so exited he spent the next 40 minutes completely engrossed in the task: first snapping his bones using his body as a lever, then sticking his fingers into the arm, pinching bundles of muscle fibers and severing them one by one, before cutting the blue arteries and the pale "noodle-like" nerves. The pain was unimportant. Only cutting through the thick white main nerve made him stop for a minute—the flood of pain, he describes, was like thrusting his entire arm "into a cauldron of magma." Finally free, he rappelled down a cliff and walked another 7 miles until he was rescued by some hikers (Ralston, 2010). How is it possible to do something so excruciatingly painful to yourself, and still manage to walk, talk, and think rationally afterwards? The answer lies within the brain, where signals from the body are interpreted. When we perceive somatosensory and nociceptive signals from the body, the experience is highly subjective and malleable by motivation, attention, emotion, and context.

The Motivation–Decision Model and Descending Modulation of Pain

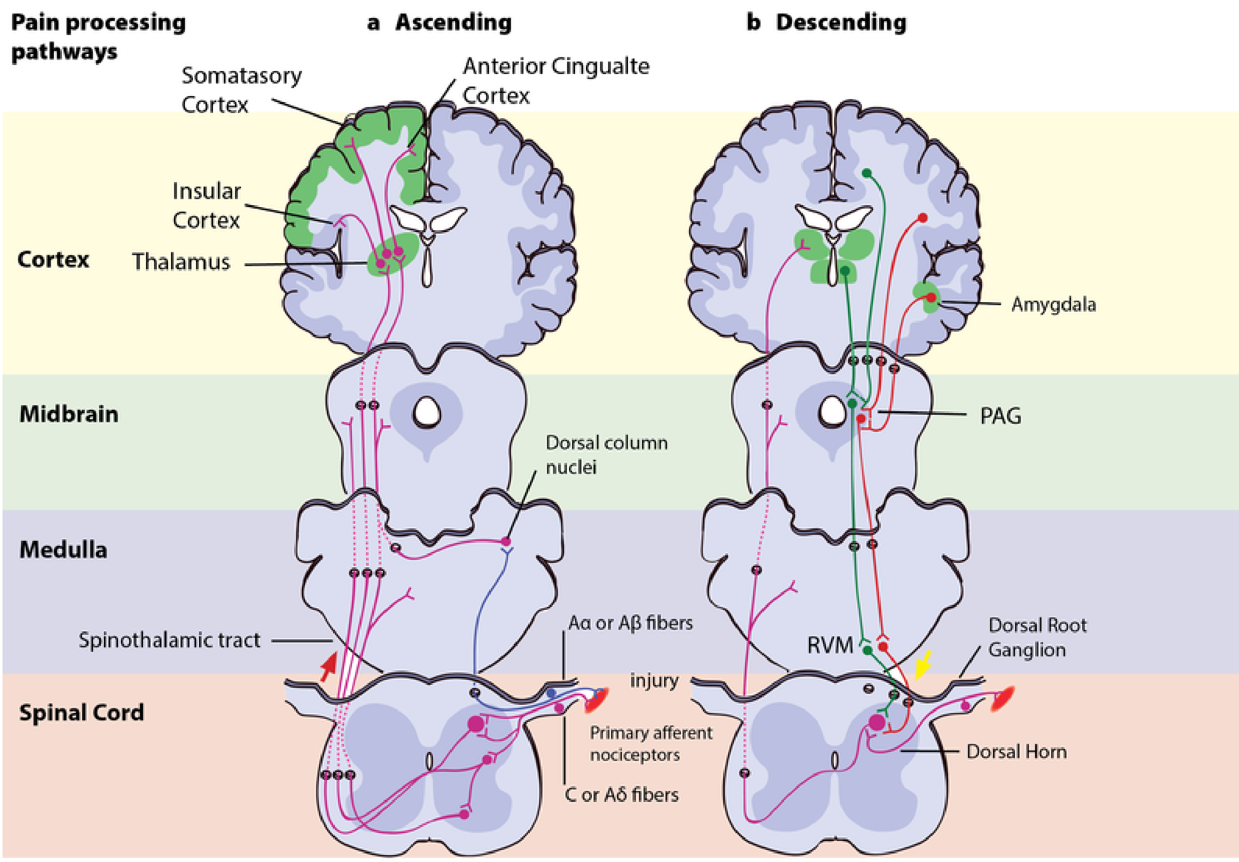
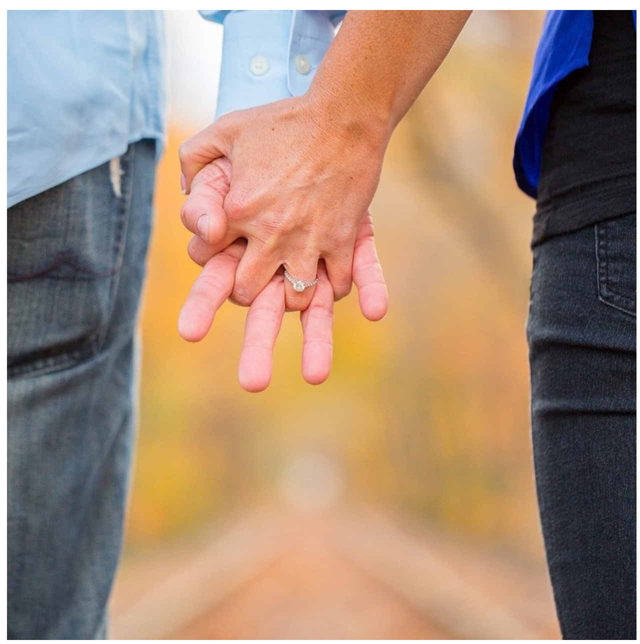


Figure 3: Pain processing pathways Left - Ascending pain pathways: An injury is signaled simultaneously via fast-conducting Aα or Aβ-fibres and slow-conducting C-pain or Aδ-fibres. The fast A-fibres signal pressure, stretching and other tissue movements to the somatosensory cortex via the dorsal column nuclei. The C-pain and Aδ-fibres sends pain information from nociceptors in the tissue or skin, and transmits these signals to second order neurons in the dorsal horn of the spinal cord. The second order neurons then cross over to the opposite side, where they form the ascending spinothalamic tract. This tract projects signals to nuclei in the medulla and midbrain on the way up to the thalamus (T). The thalamus relays the information to the somatosensory and insular cortex, as well as cortical regions mediating different aspects of the pain experience such as affective responses in the cingulate cortex.
Right - Descending pain modulation pathways: Information from the environment and certain motivational states can activate this top–down pathway. Several areas in the limbic forebrain including the anterior cingulate and insular cortex, nuclei in the amygdala and the hypothalamus (H), project to the midbrain periaqueductal grey (PAG), which then modulates ascending pain transmission from the afferent pain system indirectly through the rostral ventromedial medulla (RVM) in the brainstem. This modulating system produces analgesia by the release of endogenous opioids, and uses ON- and OFF-cells to exert either inhibitory (green) or facilitatory (red) control of nociceptive signals at the spinal dorsal horn.

According to *the motivation–decision model*, the brain automatically and continuously evaluates the pros and cons of any situation—weighing impending threats and available rewards (Fields, 2004, 2006). Anything more important for survival than avoiding the pain activates the brain's descending pain modulatory system—a top-down system involving several parts of the brain and brainstem, which inhibits nociceptive signaling so that the more important actions can be attended to (Figure 3b). In Aron's extreme case, his actions were likely based on such an unconscious decision process—taking into account his homeostatic state (his hunger, thirst, the inflammation and decay of his crushed hand slowly affecting the rest of his body), the sensory input available (the sweet smell of his dissolving skin, the silence around him indicating his solitude), and his knowledge about the threats facing him (death, or excruciating pain that won't kill him) versus the potential rewards (survival, seeing his family again). Aron's story illustrates the evolutionary advantage to being able to shut off pain: The descending pain modulatory system allows us to go through with potentially life-saving actions. However, when one has reached safety or obtained the reward, healing is more important. The very same descending system can then "crank up" nociception from the body to promote healing and motivate us to avoid potentially painful actions. To facilitate or inhibit nociceptive signals from the body, the descending pain modulatory system uses a set of ON- or OFF-cells in the brainstem, which regulates how much of the nociceptive signal reaches the brain. The descending system is dependent on opioid signaling, and analgesics like morphine relieve pain via this circuit (Petrovic, Kalso, Petersson, & Ingvar, 2002).

The Analgesic Power of Reward



Social rewards, like holding the hands or just seeing the picture of a loved one, can reduce sensations of pain. [Image: rogerl01, CC0 Public Domain, https://goo.gl/m25gce]

Thinking about the good things, like his loved ones and the life ahead of him, was probably pivotal to Aron's survival. The promise of a reward can be enough to relieve pain. Expecting pain relief (getting less pain is often the best possible outcome if you're in pain, i.e., it is a reward) from a medical treatment contributes to the placebo effect—where pain relief is due at least partly to your brain's descending modulation circuit, and such relief depends on the brain's own opioid system (Eippert et al., 2009; Eippert, Finsterbusch, Bingel, & Buchel, 2009; Levine, Gordon, & Fields, 1978). Eating tasty food, listening to good music, or feeling pleasant touch on your skin also decreases pain in both animals and humans, presumably through the same mechanism in the brain (Leknes & Tracey, 2008). In a now classic experiment, Dum and Herz

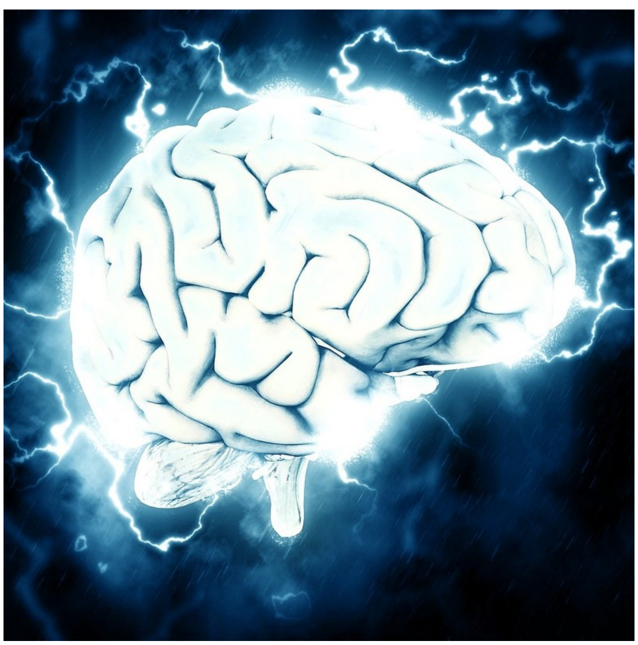
(1984) either fed rats normal rat food or let them feast on highly rewarding chocolate-covered candy (rats love sweets) while standing on a metal plate until they learned exactly what to expect when placed there. When the plate was heated up to a noxious/painful level, the rats that expected candy endured the temperature for twice as long as the rats expecting normal chow. Moreover, this effect was completely abolished when the rats' opioid (endorphin) system was blocked with a drug, indicating that the analgesic effect of reward anticipation was caused by **endorphin** release.

For Aron the climber, both the stress from knowing that death was impending and the anticipation of the reward it would be to survive probably flooded his brain with endorphins, contributing to the wave of excitement and euphoria he experienced while he carried out the amputation "like a five-year-old unleashed on his Christmas presents" (Ralston, 2010). This altered his experience of the pain from the extreme tissue damage he was causing and enabled him to focus on freeing himself. Our brain, it turns out, can modulate the perception of how unpleasant pain is, while still retaining the ability to experience the intensity of the sensation (Rainville, Duncan, Price, Carrier, & Bushnell, 1997; Rainville, Feine, Bushnell, & Duncan, 1992). Social rewards, like holding the hand of your boyfriend or girlfriend, have pain-reducing effects. Even looking at a picture of him/her can have similar effects—in fact, seeing a picture of a person we feel close to not only reduces subjective pain ratings, but also the activity in pain-related brain areas (Eisenberger et al., 2011). The most common things to do when wanting to help someone through a painful experience—being present and holding the person's hand—thus seems to have a measurably positive effect.

When Touch Becomes Painful or Pain Becomes Chronic

Chances are you've been sunburned a few times in your life and have experienced how even the lightest pat on the back or the softest clothes can feel painful on your over-sensitive skin. This condition, where innocuous touch gives a burning, tender sensation, is similar to a chronic condition called **allodynia**—where neuronal disease or injury makes touch that is normally pleasant feel unpleasantly painful. In allodynia, neuronal injury in the spinal dorsal horn causes Aβ-afferents, which are activated by non-nociceptive touch, to access nociceptive pathways (Liljencrantz et al., 2013). The result is that even gentle touch is interpreted by the brain as painful. While an acute pain response to **noxious stimuli** has a vital protective function, allodynia and other **chronic pain** conditions constitute a tremendous source of unnecessary suffering that affects millions of people. Approximately 100 million Americans suffer from chronic pain, and annual economic cost associated is estimated to be $560–$635 billion (Committee on Advancing Pain Research, Care, & Institute of Medicine, 2011). Chronic pain conditions are highly diverse, and they can involve changes on peripheral, spinal, central, and psychological levels. The mechanisms are far from fully understood, and developing appropriate treatment remains a huge challenge for pain researchers.

Chronic pain conditions often begin with an injury to a peripheral nerve or the tissue surrounding it, releasing hormones and inflammatory molecules that *sensitize* nociceptors. This makes the nerve and neighboring afferents more excitable, so that also uninjured nerves become hyperexcitable and contribute to the persistence of pain. An injury might also make neurons fire nonstop regardless of external stimuli, providing near-constant input to the pain system. **Sensitization** can also happen in the brain and in the descending modulatory system of the brainstem (Zambreanu, Wise, Brooks, Iannetti, & Tracey, 2005). Exactly on which levels the pain perception is altered in chronic pain patients can be extremely difficult to pinpoint, making treatment an often exhausting process of trial and error. Suffering from chronic pain has dramatic impacts on the lives of the afflicted. Being in pain over a longer time can lead to depression, anxiety (fear or anticipation of future pain), and immobilization, all of which may in turn exacerbate pain (Wiech & Tracey, 2009). Negative emotion and attention to pain can increase sensitization to pain, possibly by keeping the descending pain modulatory system in facilitation mode. Distraction is therefore a commonly used technique in hospitals where patients have to undergo painful treatments like changing bandages on large burns. For chronic pain patients, however, diverting attention is not a long-term solution. Positive factors like social support can reduce the risk of chronic pain after an injury, and so they can help to adjust to bodily change as a result of injury. We haveve already talked about how having a hand to hold might alleviate suffering. Chronic pain treatment should target these emotional and social factors as well as the physiological.



How powerful is the mind really? Well, the next time you have dental work done, know that some people opt for no Novocain, using only the power of their minds to overcome the pain. [Image: HypnoArt, CC0 Public Domain, https://goo.gl/m25gce]

The Power of the Mind

The context of pain and touch has a great impact on how we interpret it. Just imagine how different it would feel to Aron if someone amputated his hand against his will and for no discernible reason. Prolonged pain from injuries can be easier to bear if the incident causing them provides a positive context—like a war wound that testifies to a soldier's courage and commitment—or **phantom pain** from a hand that was cut off to enable life to carry on. The relative meaning of pain is illustrated by a recent experiment, where the same moderately painful heat was administered to participants in two different contexts—one control context where the alternative was a nonpainful heat; and another where the alternative was an intensely painful heat. In the control context, where the moderate heat was the least preferable outcome, it was (unsurprisingly) rated as painful. In the other context it was the best possible outcome, and here the exact same moderately painful heat was actually rated as *pleasant*—because it meant the intensely painful heat had been avoided. This somewhat surprising change in perception—where pain becomes pleasant because it represents relief from something worse—highlights the importance of the meaning individuals ascribe to their pain, which can have decisive effects in pain treatment (Leknes et al., 2013). In the case of touch, knowing who or what is stroking your skin can make all the difference—try thinking about slugs the next time someone strokes your skin if you want an illustration of this point. In a recent study, a group of heterosexual males were told that they were about to receive sensual caresses on the leg by either a male experimenter or by an attractive female experimenter (Gazzola et al., 2012). The study participants could not see who was touching them. Although it was always the female experimenter who performed the caress, the heterosexual males rated the otherwise pleasant sensual caresses as clearly unpleasant when they believed the male experimenter did it. Moreover, brain responses to the "male touch" in somatosensory cortex were reduced, exemplifying how top-down regulation of touch resembles top-down pain inhibition.

Pain and pleasure not only share modulatory systems—another common attribute is that we don't need to be on the receiving end of it ourselves in order to experience it. How did you feel when you read about Aron cutting through his own tissue, or "Thomas" destroying his own bones unknowingly? Did you cringe? It's quite likely that some of your brain areas processing affective aspects of pain were active even though the nociceptors in your skin and deep tissue were not firing. Pain can be experienced vicariously, as can itch, pleasurable touch, and other sensations. Tania Singer and her colleagues found in an fMRI study that some of the same brain areas that were active when participants felt pain on their own skin (anterior cingulate and insula) were also active when they were given a signal that a loved one was feeling the pain. Those who were most "empathetic" also showed the largest brain responses (Singer et al., 2004). A similar effect has been found for pleasurable touch: The posterior insula of participants watching videos of someone else's arm being gently stroked shows the same activation as if they were receiving the touch themselves (Morrison, Bjornsdotter, & Olausson, 2011).

Summary

Sensory experiences connect us to the people around us, to the rest of the world, and to our own bodies. Pleasant or unpleasant, they're part of being human. In this module, we have seen how being able to inhibit pain responses is central to our survival—and in cases like that of climber Aron Ralston, that ability can allow us to do extreme things. We have also seen how important the ability to feel pain is to our health—illustrated by young "Thomas," who keeps injuring himself because he simply doesn't notice pain. While "Thomas" has to learn to avoid harmful activities without the sensory input that normally guides us, G. L. has had to learn how to keep approaching and move about in a world she can hardly feel at all, with a body that is practically disconnected from her awareness. Too little sensation or too much of it leads to no good, no matter how pleasant or unpleasant the sensation usually feels. As long as we have nervous systems that function normally, we are able to adjust the volume of the sensory signals and our behavioral reactions according to the context we're in. When it comes to sensory signals like touch and pain, we are interpreters, not measuring instruments. The quest for understanding how our sensory–processing mechanisms can be modulated, psychologically and physiologically, promises to help researchers and clinicians find new ways to alleviate distress from chronic pain.

Outside Resources

Book: Butler, D. S., Moseley, G. L., & Sunyata. (2003). Explain pain (p. 19). Australia: Noigroup.

Book: Kringelbach, M. L., & Berridge, K. C. (Eds.). (2010). Pleasures of the brain (p. 343). Oxford, UK: Oxford University Press.

Book: Ralston, A. (2004). Between a rock and a hard place: The basis of the motion picture 127 Hours. New York, NY: Atria.

Book: Sacks, O. (1998). The man who mistook his wife for a hat: And other clinical tales. New York, NY: Simon & Schuster.

Video: BBC Documentary series "Human Senses," Episode 3: Touch and Vision
http://watchdocumentary.org/watch/human-senses-episode-03-touch-and-vision-video_f3e33c14a.html

Video: BBC Documentary "Pleasure and Pain with Michael Mosley"
http://www.bbc.co.uk/programmes/b00y377q

Video: TEDxAdelaide - Lorimer Moseley – "Why Things Hurt"
http://www.youtube.com/watch?v=gwd-wLdIHjs

Video: Trailer for the film 127 Hours, directed by Danny Boyle and released in 2010
http://www.youtube.com/watch?v=OlhLOWTnVoQ

Web: Homepage for the International Association for the Study of Pain
http://www.iasp-pain.org

Web: Proceedings of the National Academy of Sciences Colloquium "The Neurobiology of Pain"
http://www.pnas.org/content/96/14.toc#COLLOQUIUM

Web: Stanford School of Medicine Pain Management Center
http://paincenter.stanford.edu/

Website resource aiming to communicate "advances and issues in the clinical sciences as they relate to the role of the brain and mind in chronic pain disorders," led by Dr. Lorimer Moseley

http://www.bodyinmind.org/

Discussion Questions

1. Your friend has had an accident and there is a chance the injury might cause pain over a prolonged period. How would you support your friend? What would you say and do to ease the pain, and why do you think it would work?
2. We have learned that touch and pain sensation in many aspects do not reflect "objectively" the outside world or the body state. Rather, these experiences are shaped by various top-down influences, and they can even occur without any peripheral activation. This is similar to the way other sensory systems work, e.g., the visual or auditory systems, and seems to reflect a general way the brain process sensory events. Why do you think the brain interprets the incoming sensory information instead of giving a one-to-one readout the way a thermometer and other measuring instruments would? Imagine you instead had "direct unbiased access" between stimuli and sensation. What would be the advantages and disadvantages of this?
3. Feelings of pain or touch are subjective—they have a particular quality that you perceive subjectively. How can we know whether the pain you feel is similar to the pain I feel? Is it possible that modern scientists can objectively measure such subjective feelings?

Vocabulary

A-fibers
Fast-conducting sensory nerves with myelinated axons. Larger diameter and thicker myelin sheaths increases conduction speed. Aβ-fibers conduct touch signals from low-threshold mechanoreceptors with a velocity of 80 m/s and a diameter of 10 μm; Aδ-fibers have a diameter of 2.5 μm and conduct cold, noxious, and thermal signals at 12 m/s. The third and fastest conducting A-fiber is the Aα, which conducts proprioceptive information with a velocity of 120 m/s and a diameter of 20 μm.

Allodynia
Pain due to a stimulus that does not normally provoke pain, e.g., when a light, stroking touch feels painful.

Analgesia
Pain relief.

C-fibers
C-fibers: Slow-conducting unmyelinated thin sensory afferents with a diameter of 1 μm and a conduction velocity of approximately 1 m/s. C-pain fibers convey noxious, thermal, and heat signals; C-tactile fibers convey gentle touch, light stroking.

Chronic pain
Persistent or recurrent pain, beyond usual course of acute illness or injury; sometimes present without observable tissue damage or clear cause.

C-pain or Aδ-fibers
C-pain fibers convey noxious, thermal, and heat signals

C-tactile fibers
C-tactile fibers convey gentle touch, light stroking

Cutaneous senses
The senses of the skin: tactile, thermal, pruritic (itchy), painful, and pleasant.

Descending pain modulatory system
A top-down pain-modulating system able to inhibit or facilitate pain. The pathway produces analgesia by the release of endogenous opioids. Several brain structures and nuclei are part

of this circuit, such as the frontal lobe areas of the anterior cingulate cortex, orbitofrontal cortex, and insular cortex; and nuclei in the amygdala and the hypothalamus, which all project to a structure in the midbrain called the periaqueductal grey (PAG). The PAG then controls ascending pain transmission from the afferent pain system indirectly through the rostral ventromedial medulla (RVM) in the brainstem, which uses ON- and OFF-cells to inhibit or facilitate nociceptive signals at the spinal dorsal horn.

Endorphin
An endogenous morphine-like peptide that binds to the opioid receptors in the brain and body; synthesized in the body's nervous system.

Exteroception
The sense of the external world, of all stimulation originating from outside our own bodies.

Interoception
The sense of the physiological state of the body. Hunger, thirst, temperature, pain, and other sensations relevant to homeostasis. Visceral input such as heart rate, blood pressure, and digestive activity give rise to an experience of the body's internal states and physiological reactions to external stimulation. This experience has been described as a representation of "the material me," and it is hypothesized to be the foundation of subjective feelings, emotion, and self-awareness.

Nociception
The neural process of encoding noxious stimuli, the sensory input from nociceptors. Not necessarily painful, and crucially not necessary for the experience of pain.

Nociceptors
High-threshold sensory receptors of the peripheral somatosensory nervous system that are capable of transducing and encoding noxious stimuli. Nociceptors send information about actual or impending tissue damage to the brain. These signals can often lead to pain, but nociception and pain are not the same.

Noxious stimulus
A stimulus that is damaging or threatens damage to normal tissues.

Ocial touch hypothesis
Proposes that social touch is a distinct domain of touch. C-tactile afferents form a special pathway that distinguishes social touch from other types of touch by selectively firing in response to touch of social-affective relevance; thus sending affective information parallel to

the discriminatory information from the Aβ-fibers. In this way, the socially relevant touch stands out from the rest as having special positive emotional value and is processed further in affect-related brain areas such as the insula.

Pain

Defined as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage,” according to the International Association for the Study of Pain.

Phantom pain

Pain that appears to originate in an amputated limb.

Placebo effect

Effects from a treatment that are not caused by the physical properties of a treatment but by the meaning ascribed to it. These effects reflect the brain’s own activation of modulatory systems, which is triggered by positive expectation or desire for a successful treatment. Placebo analgesia is the most well-studied placebo effect and has been shown to depend, to a large degree, on opioid mechanisms. Placebo analgesia can be reversed by the pharmacological blocking of opioid receptors. The word “placebo” is probably derived from the Latin word “placebit” (“it will please”).

Sensitization

Increased responsiveness of nociceptive neurons to their normal input and/or recruitment of a response to normally subthreshold inputs. Clinically, sensitization may only be inferred indirectly from phenomena such as hyperalgesia or allodynia. Sensitization can occur in the central nervous system (central sensitization) or in the periphery (peripheral sensitization).

Social touch hypothesis

Proposes that social touch is a distinct domain of touch. C-tactile afferents form a special pathway that distinguishes social touch from other types of touch by selectively firing in response to touch of social-affective relevance; thus sending affective information parallel to the discriminatory information from the Aβ-fibers. In this way, the socially relevant touch stands out from the rest as having special positive emotional value and is processed further in affect-related brain areas such as the insula.

Somatosensory cortex

Consists of primary sensory cortex (S1) in the postcentral gyrus in the parietal lobes and secondary somatosensory cortex (S2), which is defined functionally and found in the upper bank of the lateral sulcus, called the parietal operculum. Somatosensory cortex also includes parts of the insular cortex.

Somatotopically organized
When the parts of the body that are represented in a particular brain region are organized topographically according to their physical location in the body (see Figure 2 illustration).

Spinothalamic tract
Runs through the spinal cord's lateral column up to the thalamus. C-fibers enter the dorsal horn of the spinal cord and form a synapse with a neuron that then crosses over to the lateral column and becomes part of the spinothalamic tract.

Transduction
The mechanisms that convert stimuli into electrical signals that can be transmitted and processed by the nervous system. Physical or chemical stimulation creates action potentials in a receptor cell in the peripheral nervous system, which is then conducted along the axon to the central nervous system.

References

Auvray, M., Myin, E., & Spence, C. (2010). The sensory-discriminative and affective-motivational aspects of pain. *Neuroscience and Biobehavioral Reviews, 34*(2), 214–223.

Committee on Advancing Pain Research, Care, & Institute of Medicine. (2011). *Relieving pain in America: A blueprint for transforming prevention, care, education, and research*. Washington, D.C.: National Academy Press.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience, 3*(8), 655–666.

Dum, J., & Herz, A. (1984). Endorphinergic modulation of neural reward systems indicated by behavioral changes. *Pharmacology Biochemistry and Behavior, 21*(2), 259–266.

Dunbar, R. I. (2010). The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. *Neuroscience & Biobehavioral Reviews, 34*(2), 260–268.

Eippert, F., Bingel, U., Schoell, E. D., Yacubian, J., Klinger, R., Lorenz, J., & Buchel, C. (2009). Activation of the opioidergic descending pain control system underlies placebo analgesia. *Neuron, 63*(4), 533–543.

Eippert, F., Finsterbusch, J., Bingel, U., & Buchel, C. (2009). Direct evidence for spinal cord involvement in placebo analgesia. *Science, 326*(5951), 404.

Eisenberger, N. I., Master, S. L., Inagaki, T. K., Taylor, S. E., Shirinyan, D., Lieberman, M. D., & Naliboff, B. D. (2011). Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proceedings of the National Academy of Sciences of the United States of America, 108*(28), 11721–11726.

Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science, 13*(2), 271–278.

Fields, H. L. (2006). A motivation-decision model of pain: The role of opioids. In H. Flor, E. Kalso, & J. O. Dostrovsky (Eds.), *Proceedings of the 11th World Congress on Pain* (pp. 449–459). Seattle, WA: IASP.

Fields, H. L. (2004). State-dependent opioid control of pain. *Nature Reviews Neuroscience, 5*(7), 565–575.

Fleming, A. S., O'Day, D. H., & Kraemer, G. W. (1999). Neurobiology of mother-infant interactions: Experience and central nervous system plasticity across development and generations. *Neuroscience and Biobehavioral Reviews, 23*(5), 673–685.

Gazzola, V., Spezio, M. L., Etzel, J. A., Castelli, F., Adolphs, R., & Keysers, C. (2012). Primary somatosensory cortex discriminates affective significance in social touch. *Proceedings of* *the National Academy of Sciences of the United States of America, 109*(25), E1657–1666.

Harlow, H. F., & Suomi, S. J. (1970). Nature of love—simplified. *American Psychologist, 25*(2), 161–168.

Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology, 14*(1–2), 155–161.

Leknes, S., & Tracey, I. (2008). A common neurobiology for pain and pleasure. *Nature Reviews Neuroscience, 9*(4), 314–320.

Leknes, S., Berna, C., Lee, M. C., Snyder, G. D., Biele, G., & Tracey, I. (2013). The importance of context: When relative relief renders pain pleasant. *Pain, 154*(3), 402–410.

Levine, J. D., Gordon, N. C., & Fields, H. L. (1978). The mechanism of placebo analgesia. *Lancet, 2*(8091), 654–657.

Liljencrantz, J., Björnsdotter, M., Morrison, I., Bergstrand, S., Ceko, M., Seminowicz, D. A., . . . Olausson, H. (2013). Altered C-tactile processing in human dynamic tactile allodynia. *Pain, 154*(2), 227–234.

Löken, L. S., Wessberg, J., Morrison, I., McGlone, F., & Olausson, H. (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nature Neuroscience, 12*(5), 547–548.

McGlone, F., & Reilly, D. (2010). The cutaneous sensory system. *Neuroscience & Biobehavioral Reviews, 34*(2), 148–159.

Minde, J., Toolanen, G., Andersson, T., Nennesmo, I., Remahl, I. N., Svensson, O., & Solders, G. (2004). Familial insensitivity to pain (HSAN V) and a mutation in the NGFB gene. A neurophysiological and pathological study. *Muscle & Nerve, 30*(6), 752–760.

Morrison, I., Bjornsdotter, M., & Olausson, H. (2011). Vicarious responses to social touch in posterior insular cortex are tuned to pleasant caressing speeds. *Journal of Neuroscience, 31*(26), 9554–9562.

Morrison, I., Löken, L. S., & Olausson, H. (2010). The skin as a social organ. *Experimental Brain Research, 204*(3), 305–314.

Olausson, H., Lamarre, Y., Backlund, H., Morin, C., Wallin, B. G., Starck, G., . . . Bushnell, M. C. (2002). Unmyelinated tactile afferents signal touch and project to insular cortex. *Nature Neuroscience, 5*(9), 900–904.

Olausson, H., Wessberg, J., Morrison, I., McGlone, F., & Vallbo, Å. (2010). The neurophysiology of unmyelinated tactile afferents. *Neuroscience & Biobehavioral Reviews, 34*(2), 185–191.

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: An integrated review of human literature. *Psychopharmacology, 214*(1), 55–70.

Petrovic, P., Kalso, E., Petersson, K. M., & Ingvar, M. (2002). Placebo and opioid analgesia—Imaging a shared neuronal network. *Science, 295*(5560), 1737–1740.

Price, D. D. (2000). Psychological and neural mechanisms of the affective dimension of pain. *Science, 288*(5472), 1769–1772.

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science, 277*(5328), 968–971.

Rainville, P., Feine, J. S., Bushnell, M. C., & Duncan, G. H. (1992). A psychophysical comparison of sensory and affective responses to four modalities of experimental pain. *Somatosensory and Motor Research, 9*(4), 265–277.

Ralston, A. (2010, December 31). Aron Ralston's 127 hours: This is going to make one hell of a story. . . Retrieved from The Telegraph: http://www.telegraph.co.uk/culture/books/8233-734/Aron-Ralstons-127-Hours-This-is-going-to-make-one-hell-of-a-story.-.-..html

Sacks, O. (1985). The disembodied lady. In *The man who mistook his wife for a hat and other clinical tales* (pp. 43–54) New York, NY: Simon & Schuster.

Simons, D. J., & Land, P. W. (1987). Early experience of tactile stimulation influences organization of somatic sensory cortex. *Nature, 326*(6114), 694–697.

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*(5661), 1157–1162.

Wiech, K., & Tracey, I. (2009). The influence of negative emotions on pain: Behavioral effects and neural mechanisms. *NeuroImage, 47*(3), 987–994.

Zambreanu, L., Wise, R. G., Brooks, J. C. W., Iannetti, G. D., & Tracey, I. (2005). A role for the brainstem in central sensitisation in humans. Evidence from functional magnetic resonance imaging. *Pain, 114*(3), 397–407.

16
The Vestibular System

Dora Angelaki & J. David Dickman

The vestibular system functions to detect head motion and position relative to gravity and is primarily involved in the fine control of visual gaze, posture, orthostasis, spatial orientation, and navigation. Vestibular signals are highly processed in many regions of the brain and are involved in many essential functions. In this module, we provide an overview of how the vestibular system works and how vestibular signals are used to guide behavior.

Learning Objectives

- Define the basic structures of the vestibular receptor system.
- Describe the neuroanatomy of the vestibuloocular, vestibulospinal, and vestibulo-thalamo-cortical pathways.
- Describe the vestibular commissural system.
- Describe the different multisensory cortical areas for motion perception.

Introduction

Remember the dizzy feeling you got as a child after you jumped off the merry-go-round or spun around like a top? These feelings result from activation of the **vestibular system**, which detects our movements through space but is not a conscious sense like vision or hearing. In fact, most vestibular functions are imperceptible, but vestibular-related sensations such as motion sickness can pop up rapidly when riding on a roller coaster, having a bumpy plane ride, or a sailing a boat in rough seas. However, these sensations are really side effects and the vestibular system is actually extremely important for everyday activities, with vestibular signals being involved in much of the brain's information processing that controls such fundamental functions as balance, posture, gaze stabilization, spatial orientation, and navigation, to name a few. In many regions of the brain, vestibular information is combined with signals from the other senses as well as with motor information to give rise to motion perception, body awareness, and behavioral control. Here, we will explore the workings of the vestibular system and consider some of the integrated computations the brain performs using vestibular signals to guide our common behavior.

Although you may be most aware of the vestibular system in situations where you experience motion sickness, at other times its working in the background to provide you with essential functions like balance and stable vision. [Image: Leo Reynolds, https://goo.gl/N8D94w, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Structure of the vestibular receptors

The vestibular receptors lie in the inner ear next to the auditory cochlea. They detect rotational motion (head turns), linear motion (translations), and tilts of the head relative to gravity and transduce these motions into neural signals that can be sent to the brain. There are five vestibular receptors in each ear (Hearing module, Figure 1- **http://noba.to/jry3cu78**), including three **semicircular canals** (horizontal, anterior, and posterior) that transduce rotational angular accelerations and two **otolith receptors** (utricle and saccule) that transduce linear accelerations (Lindeman, 1969). Together, the semicircular canals and otolith organs can respond to head motion and maintained static head position relative to gravity in all directions in 3D space.

These receptors are contained in a series of interconnected fluid filled tubes that are protected by a dense, overlying bone (Iurato, 1967). Each of the three semicircular canals lies in a plane that is orthogonal to the other two. The horizontal semicircular canal lies in a roughly horizontal head plane, whereas the anterior and posterior semicircular canals lie vertically in the head (Blanks, Curthoys, Bennett, & Markham, 1985). The semicircular canal receptor cells, termed **hair cells**, are located only in the middle of the circular tubes in a special epithelium, covered by a gelatinous membrane that stretches across the tube to form a fluid-tight seal like the skin of a drum (Figures 1A and 1B). Hair cells are so named due to an array of nearly 100 staggered-height **stereocilia** (like a church pipe organ) that protrude from the top of the cell into the overlying gelatin membrane (Wersäll, 1956). The shortest stereocilia are at one end of the cell and the tallest at the other (Lindeman, 1969). When the head is rotated, the fluid in the semicircular canals lags behind the head motion and pushes on the gelatin membrane, which bends the stereocilia.

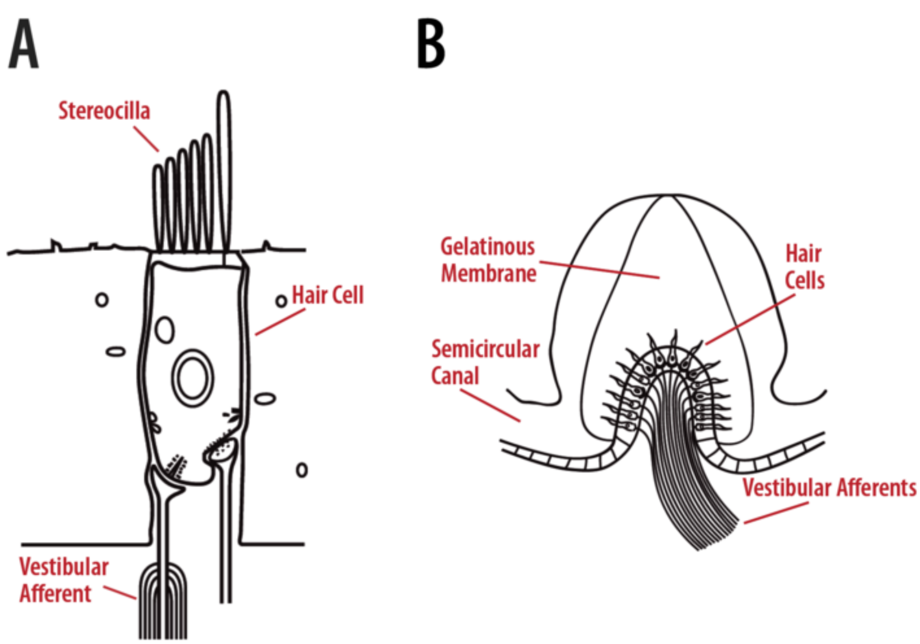


Figure 1. Receptor hair cells and semicircular canal structure. A) Vestibular hair cell in the receptor epithelium with stereocilia in the apical surface of the cell. Innervating afferent and efferent neurons make synaptic contacts with the basal surface of the cell. B) Semicircular canal structure showing the fluid duct, the hair cell stereocilia embedded in a gelatinous membrane on top of the hair cells, and the innervating afferent fibers.

As shown in Figure 2, when the head moves toward the receptor hair cells (e.g., left head turns for the left horizontal semicircular canal), the stereocilia are bent toward the tallest end and special mechanically gated ion channels in the tips of the cilia open, which excites (depolarizes) the cell (Shotwell, Jacobs, & Hudspeth, 1981). Head motion in the opposite direction causes bending toward the smallest stereocilia, which closes the channels and inhibits (hyperpolarizes) the cell. The left and right ear semicircular canals have opposite polarity, so for example, when you turn your head to the left, the receptors in the left horizontal semicircular canal will be excited while right ear horizontal canal receptors will be inhibited (Figure 3). The same relationship is true for the vertical semicircular canals. Vestibular afferent nerve fibers innervate the base of the hair cell and increase or decrease their neural firing rate as the receptor cell is excited or inhibited (Dickman and Correia, 1989), respectively, and then carry these signals regarding head rotational motion to the brain as part of the vestibulocochlear nerve (Cranial nerve VIII). They enter the brainstem and terminate in the ipsilateral vestibular nuclei, cerebellum, and reticular formation (Carleton & Carpenter, 1984; Dickman & Fang, 1996). The primary vestibular hair cell and afferent neurotransmitters are glutamate and aspartate. Due to the mechanical properties of the vestibular receptor system,

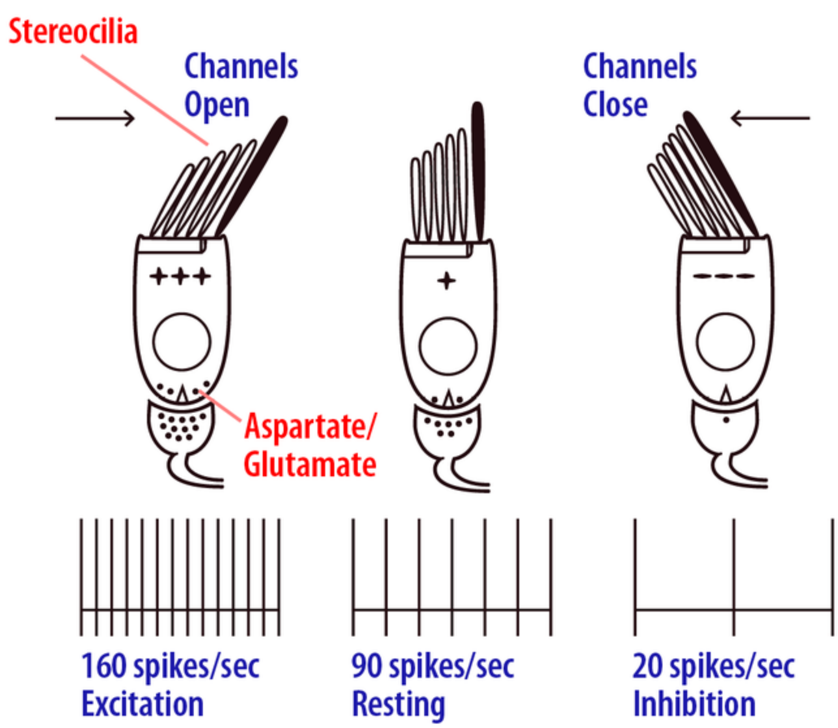


Figure 2. Vestibular receptor cell directional selectivity. Middle) At rest, hair cells release some neurotransmitter, producing a high, spontaneous firing rate in the innervating afferent fibers. Left) When the stereocilia are displaced toward the kinocilium, the cell is depolarized and the afferent firing rate is increased. Right) When the stereocilia are displaced away from the kinocilium, the cell is hyperpolarized and the afferent firing rate is decreased

rotational accelerations of the head are integrated into velocity signals (Van Egmond, Groen, & Jongkess, 1949) that are then encoded by semicircular canal afferents (Fernandez & Goldberg, 1971). Detection thresholds for rotational motion have shown that afferents can discriminate differences in head velocity on the order of 2 deg/sec, but also are sensitive to a broad range of natural head movements up to high head speeds in the hundreds of deg/sec (as you might experience when you make a fast head turn toward a loud sound, or are performing gymnastics; Sadeghi, Chacron, Taylor, & Cullen, 2007; Yu, Dickman, & Angelaki, 2012).

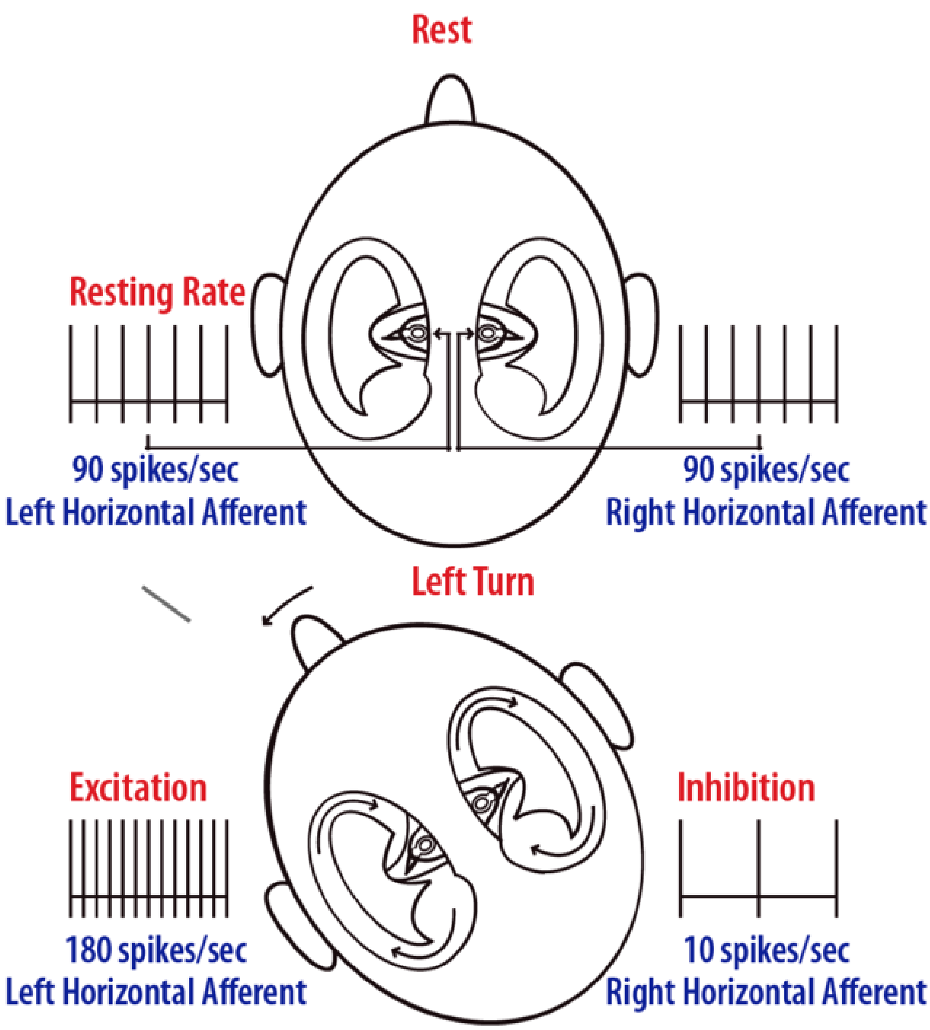


Figure 3. Receptor hair cells in the otolith organs. Receptor cells have stereocilia embedded in the gelatinous membrane, which is covered by thousands of calcium carbonite otoconia. Receptor cells are polarized in opposite directions relative to a central location and are innervated by VIIIth nerve afferent fibers.

Otolith receptors are sensitive to linear accelerations and tilts of the head relative to gravity (Fernandez & Goldberg, 1976a). The utricle otolith receptor lies parallel to the horizontal semicircular canal and the saccule receptor lies vertical in the head (Hearing module, Figure 1- **http://noba.to/jry3cu78**). As shown in Figure 4, a special otolith epithelium contains receptor hair cells whose stereocilia extend into a gelatin membrane that is covered by a layer of calcium carbonate crystals, termed **otoconia**, like rocks piled up to form a jetty (Lindeman, 1969). Otoconia are not affected by fluid movements but instead are displaced by linear accelerations, including translations (e.g., forward/backward or upward/downward motions) or changes in head position relative to gravity. These linear accelerations produce displacements of the otoconia (due to their high mass), much like rocks rolling down a hill or your coffee cup falling off the car dashboard when you push the gas pedal. Movements of the otoconia bend the hair cell stereocilia and open/close channels in a similar way to that described for the semicircular canals. However, otolith hair cells are polarized such that the tallest stereocilia are pointing toward the center of the utricle and away from the center in the saccule, which effectively splits the receptors into two opposing groups (Flock, 1964; Lindeman, 1969). In this way, some hair cells are excited and some inhibited for each linear motion force or head tilt experienced, with the population of receptors and their innervating afferents being **directionally tuned** to all motions or head tilts in 3D space (Fernandez & Goldberg, 1976b).

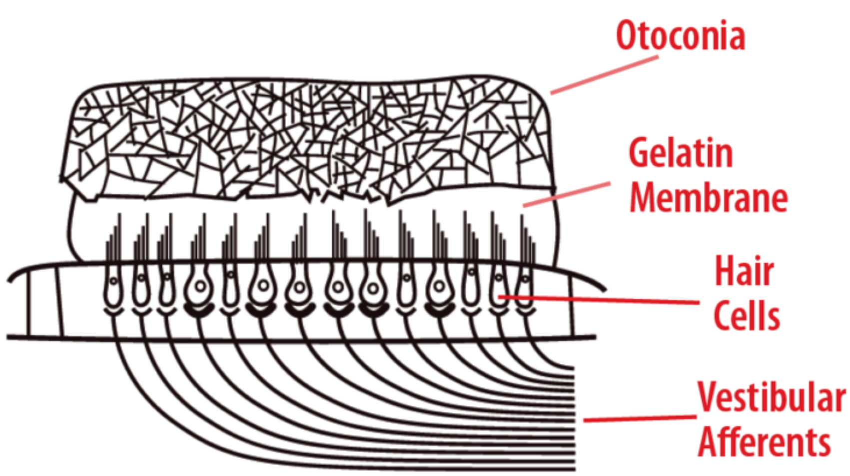


Figure 4. Vestibular response to horizontal plane head rotation. A) When the head is stationary, afferent fibers on both the sides of the head have equivalent firing so there is no sense of motion. B) When the head turns to the left, all of the left horizontal semicircular canal hair cells are excited and afferent fibers increase their firing rate. Conversely, right horizontal canal afferents decrease their firing rate

All vestibular hair cells and afferents receive connections from **vestibular efferents**, which are fibers projecting from the brain out to the vestibular receptor organs, whose function is not well understood. It is thought that efferents control the sensitivity of the receptor (Boyle, Carey, & Highstein, 1991). The primary efferents neurotransmitter is **acetylcholine** (Anniko & Arnold, 1991).

The vestibular nuclei

The vestibular nuclei comprise a large set of neural elements in the brainstem that receive motion and other multisensory signals, then regulate movement responses and sensory experience. Many vestibular nuclei neurons have reciprocal connections with the cerebellum that form important regulatory mechanisms for the control of eye movements, head movements, and posture. There are four major vestibular nuclei that lie in the rostral medulla and caudal pons of the brainstem; all receive direct input from vestibular afferents (Brodal, 1984; Precht & Shimazu, 1965). Many of these nuclei neurons receive convergent motion information from the opposite ear through an inhibitory commissural pathway that uses **gamma-aminobutyric acid** (GABA) as a neurotransmitter (Kasahara & Uchino, 1974; Shimazu & Precht, 1966). The commissural pathway is highly organized such that cells receiving horizontal excitatory canal signals from the ipsilateral ear will also receive contralateral inhibitory horizontal canal signals from the opposite ear This fact gives rise to a "push-pull" vestibular function, whereby directional sensitivity to head movement is coded by opposing receptor signals. Because vestibular nuclei neurons receive information from bilateral inner ear receptors and because they maintain a high spontaneous firing rate (nearly 100 impulses/sec), they are thought to act to "compare" the relative discharge rates of left vs. right canal afferent firing activity. For example, during a leftward head turn, left brainstem nuclei neurons receive high firing-rate information from the left horizontal canal and low firing-rate information from the right horizontal canal. The comparison of activity is interpreted as a left head turn. Similar nuclei neuron responses exist when the head is pitched or rolled, with the vertical semicircular canals being stimulated by the rotational motion in their sensitivity planes. However, the opposing push-pull response from the vertical canals occurs with the anterior semicircular canal in one ear and the co-planar posterior semicircular canal of the opposite ear. Damage or disease that interrupts inner ear signal information from one side of the head can change the normal resting activity in the VIIIth nerve afferent fibers and will be interpreted by the brain as a head rotation, even though the head is stationary. These effects often lead to illusions of spinning or rotating that can be quite upsetting and may produce nausea or vomiting. However, over time the commissural fibers provide for **vestibular compensation**, a process by which the loss of unilateral vestibular receptor function is partially restored centrally and behavioral responses, such as the **vestibuloocular**

reflex (VOR) and postural responses, mostly recover (Beraneck et al., 2003; Fetter & Zee, 1988,; Newlands, Hesse, Haque, & Angelaki, 2001; Newlands & Perachio, 1990).

In addition to the commissural pathway, many vestibular nuclei neurons receive **proprioceptive** signals from the spinal cord regarding muscle movement and position, visual signals regarding spatial motion, other multisensory (e.g., trigeminal) signals, and higher order signals from the cortex. It is thought that the cortical inputs regulate fine gaze and postural control, as well as suppress the normal **compensatory reflexes** during motion in order to elicit volitional movements. Of special significance are convergent signals from the semicircular canal and otolith afferents that allow central vestibular neurons to compute specific properties of head motion (Dickman & Angelaki, 2002). For example, Einstein (1907) showed that linear accelerations are equivalent whether they arise from translational motion or from tilts of the head relative to gravity. The otolith receptors cannot discriminate between the two, so how is it that we can tell the difference between when we are translating forward and tilting backward, where the linear acceleration signaled by the otolith afferents is the same? Vestibular nuclei and cerebellar neurons use convergent signals from both the semicircular canals and the otolith receptors to discriminate between tilt and translation, and as a result, some cells encode head tilt (Zhou, 2006) while other cells encode translational motion (Angelaki, Shaikh, Green, & Dickman, 2004).

Vestibuloocular system

The vestibular system is responsible for controlling **gaze stability** during motion (Crane & Demer, 1997). For example, if we want to read the sign in a store window while walking by, we must maintain foveal fixation on the words while compensating for the combined rotational and translational head movements incurred during our stride. The vestibular system regulates compensatory eye, neck, spinal, and limb movements in order to maintain gaze (Keshner & Peterson, 1995). One of the major components contributing to gaze stability is the VOR, which produces reflexive eye movements that are equal in magnitude and opposite in direction to the perceived head motion in 3D space (Wilson et al., 1995). The VOR is so accurate and fast that it allows people to maintain visual fixation on objects of interest while experiencing demanding motion conditions, such as running, skiing, playing tennis, and driving. In fact, gaze stabilization in humans has been shown to be completely compensatory (essentially perfect) for most natural behaviors. To produce the VOR, vestibular neurons must control each of the six pairs of eye muscles in unison through a specific set of connections to the **oculomotor nuclei** (Ezure & Graf, 1984). The anterior and posterior semicircular canals along with the saccule control vertical and torsional (turning of the eye around the line of sight) eye movements, while the horizontal canals and the utricle control horizontal eye movements.

To understand how the VOR works, let's take the example of the compensatory response for a leftward head turn while reading the words on a computer screen. The basic pathway consists of horizontal semicircular canal afferents that project to specific neurons in the vestibular nuclei. These nuclei cells, in turn, send an excitatory signal to the contralateral abducens nucleus, which projects through the sixth cranial nerve to innervate the lateral rectus muscle (Figure 5). Some abducens neurons send an excitatory projection back across

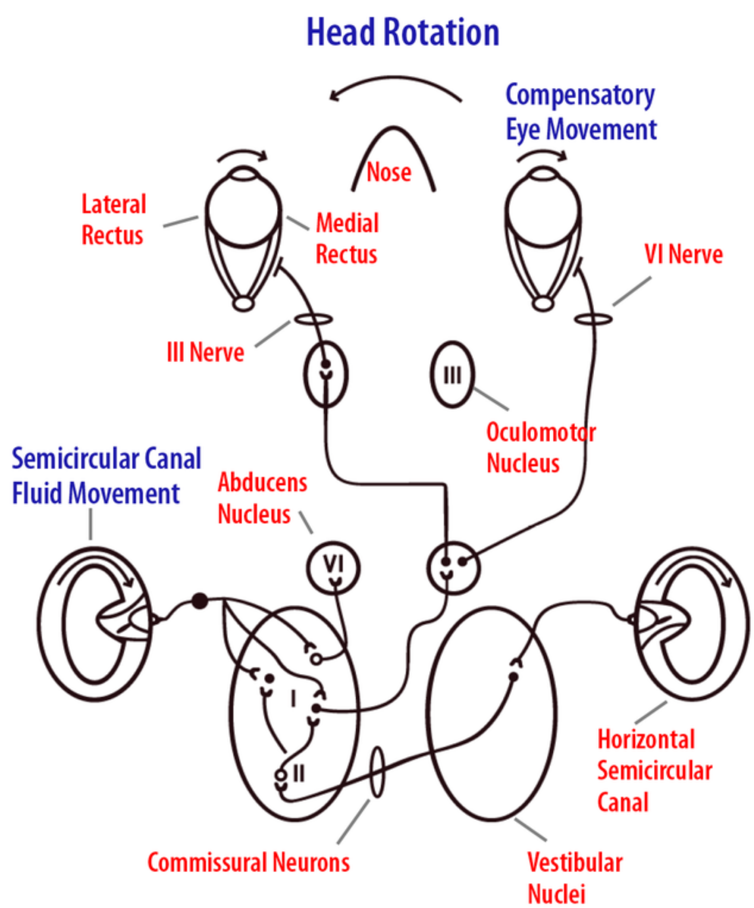


Figure 5. Vestibuloocular reflex. During a leftward head turn, the left horizontal semicircular canal receptors are excited, while the right ear receptors are inhibited. The left excitatory signals excite vestibular nuclei neurons. These cells project across the brain to excite motor neurons in the right abducens nucleus (VI) that excite the lateral rectus muscle of the right eye and to cells in the oculomotor nucleus (III) that excite the medial rectus muscle of the left eye. This moves both eyes to the right to exactly match the leftward head movement and stabilize visual gaze upon a target of interest. The right ear inhibitory signals cross to neurons in the left vestibular nucleus that decrease their firing rate. These cells are inhibitory and decrease their firing rate to further increase the response of rightward motor eye muscle cells.

the midline to a subdivision of cells in the ipsilateral oculomotor nucleus, which, in turn, projects through the third cranial nerve to innervate the right (ipsilateral) medial rectus muscle. When a leftward head turn is made, the left horizontal canal vestibular afferents will increase their firing rate and consequently increase the activity of vestibular nuclei neurons projecting to the opposite (contralateral) right abducens nucleus. The abducens neurons produce contraction of the right lateral rectus and, through a separate cell projection to the left oculomotor nucleus, excite the left medial rectus muscles. In addition, matching bilateral inhibitory connections relax the left lateral rectus and right medial rectus eye muscles. The resulting rightward eye movement for both eyes stabilizes the object of interest upon the retina for greatest visual acuity.

During linear translations, a different type of VOR also occurs (Paige & Tomko, 1991). For example, sideways motion to the left results in a horizontal rightward eye movement to maintain visual stability on an object of interest. In a similar manner, vertical up–down head movements (such as occur while walking or running) elicit oppositely directed vertical eye movements (Angelaki, McHenry, & Hess, 2000). For these reflexes, the amplitude of the translational VOR depends on viewing distance. This is due to the fact that the vergence angle (i.e., the angle between the lines of sight for each eye) varies as a function of the inverse of the distance to the viewed visual object (Schwarz, Busettini, & Miles, 1989). Visual objects that are far away (2 meters or more) require no vergence angle, but as the visual objects get closer (e.g., when holding your finger close to your nose), a large vergence angle is needed. During translational motion, the eyes will change their vergence angle as the visual object moves from close to farther away (or vice versa). These responses are a result of activation of the otolith receptors, with connections to the oculomotor nuclei similar to those described above for the rotational vestibuloocular reflex. With tilts of the head, the resulting eye movement is termed torsion, and consists of a rotational eye movement around the line of sight that is in the direction opposite to the head tilt. As mentioned above, there are major reciprocal connections between the vestibular nuclei and the cerebellum. It has been well established that these connections are crucial for adaptive motor learning in the vestibuloocular reflex (Lisberger, Pavelko, & Broussard, 1994).

Vestibulo-spinal network

There are two vestibular descending pathways that regulate body muscle responses to motion and gravity, consisting of the lateral vestibulo-spinal tract (LVST) and the medial vestibulo-spinal tract (MVST). Reflexive control of head and neck muscles arises through the neurons in the medial vestibulospinal tract (MVST). These neurons comprise the rapid vestibulocollic reflex (VCR) that serves to stabilize the head in space and participates in gaze control (Peterson, Goldber, Bilotto, & Fuller, 1985). The MVST neurons receive input from vestibular receptors and the cerebellum, and somatosensory information from the spinal cord. MVST neurons carry both excitatory and inhibitory signals to innervate neck flexor and extensor motor neurons in the spinal cord. For example, if one trips over a crack in the pavement while walking, MVST neurons will receive downward and forward linear acceleration signals from the otolith receptors and forward rotation acceleration signals from the vertical semicircular canals. The VCR will compensate by providing excitatory signals to the dorsal neck flexor muscles and inhibitory signals to the ventral neck extensor muscles, which moves the head upward and opposite to the falling motion to protect it from impact.

Similar to how the gyroscope in your cell phone can detect which angle it's at to adjust the screen, so do the MSVT neurons inform your body which direction it's headed in order to make the appropriate bodily adjustments. [Image: Kleman Gellek, https://goo.gl/DR9rpR, CC BY 4.0, https://goo.gl/QuGXFp]

The LVST comprises a topographic organization of vestibular nuclei cells that receive substantial input from the cerebellum, proprioceptive inputs from the spinal cord, and convergent afferent signals from vestibular receptors. LVST fibers project ipsilateral to many levels of motor neurons in the cord to provide coordination of different muscle groups for postural control (Shinoda, Sugiuchi, Futami, Ando, & Kawasaki, 1994). LVST neurons contain either acetylcholine or glutamate as a neurotransmitter and exert an excitatory influence upon extensor muscle motor neurons. For example, LVST fibers produce extension of the contralateral axial and limb musculature when the body is tilted sideways. These actions serve to stabilize the body's center of gravity in order to preserve upright posture.

Vestibulo-autonomic control

Some vestibular nucleus neurons send projections to the reticular formation, dorsal pontine nuclei, and nucleus of the solitary tract. These connections regulate breathing and circulation through compensatory vestibular autonomic responses that stabilize respiration and blood pressure during body motion and changes relative to gravity. They may also be important for induction of motion sickness and emesis.

Vestibular signals in the thalamus and cortex

The cognitive perception of motion, spatial orientation, and navigation through space arises through multisensory information from vestibular, visual, and somatosensory signals in the thalamus and cortex (Figure 6A). Vestibular nuclei neurons project bilaterally to the several thalamic regions. Neurons in the ventral posterior group respond to either vestibular signals alone, or to vestibular plus somatosensory signals, and projects to primary somatosensory cortex (area 3a, 2v), somatosensory association cortex, posterior parietal cortex (areas 5 and 7), and the insula of the temporal cortex (Marlinski & McCrea, 2008; Meng, May, Dickman, & Angelaki, 2007). The posterior nuclear group (PO), near the medial geniculate body, receives both vestibular and auditory signals as well as inputs from the superior colliculus and spinal cord, indicating an integration of multiple sensory signals. Some anterior pulvinar neurons also respond to motion stimuli and project to cortical area 3a, the posterior insula, and the temporo-parietal cortex (PIVC). In humans, electrical stimulation of the thalamic areas produces sensations of movement and sometimes dizziness.

Area 2v cells respond to motion, and electrical stimulation of this area in humans produces sensations of moving, spinning, or dizziness. Area 3a lies at the base of the central sulcus adjacent to the motor cortex and is thought to be involved in integrative motor control of the head and body (Guldin, Akbarian, & Grusser, 1992). Neurons in the PIVC are multisensory, responding to body motion, somatosensory, proprioceptive, and visual motion stimuli (Chen, DeAngelis, & Angelaki, 2011; Grusser, Pause, & Schreiter, 1982). PIVC and areas 3a and 2v are heavily interconnected. Vestibular neurons also have been observed in the posterior parietal cortex; in area 7, in the ventral intraparietal area (VIP), the medial intraparietal area (MIP), and the medial superior temporal area (MST). VIP contains multimodal neurons involved in spatial coding. MIP and MST neurons respond to body motion through space by multisensory integration of visual motion and vestibular signals (Gu, DeAngelis, & Angelaki , 2007) and many MST cells are directly involved in heading perception (Gu, Watkins, Angelaki, & DeAngelis, 2006). Lesions of the parietal cortical areas can result in confusions in spatial awareness. Finally, areas involved with the control of saccades and pursuit eye movements, including area 6, area 8, and the superior frontal gyrus, receive vestibular signals (Fukushima, Sato, Fukushima, Shinmei, & Kaneko, 2000). How these different cortical regions contribute to our perception of motion and spatial orientation is still not well understood.

Spatial orientation and navigation

Our ability to know where we are and to navigate different spatial locations is essential for survival. It is believed that a cognitive map of our environment is created through exploration

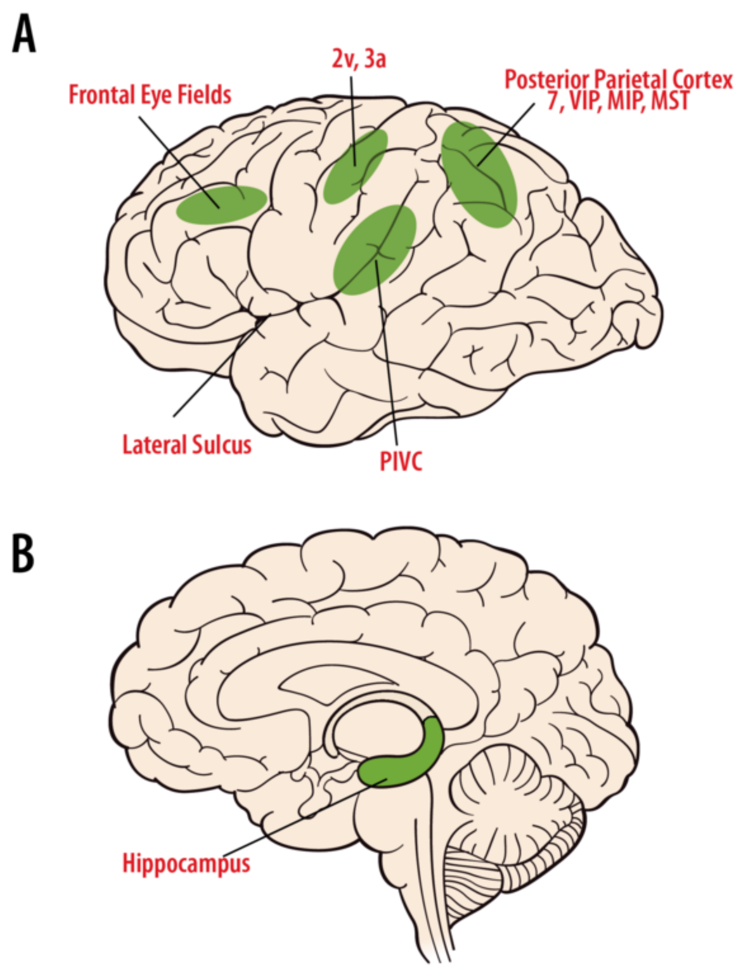


Figure 6. Cortical regions of the brain known to be involved with vestibular processing. A) The frontal eye fields control eye movements and receive vestibular motion information. Areas 2v and 3a are somatosensory areas that map body location and movement signals. Area PIVC responds to body and head motion information. The posterior parietal cortex is involved with motion perception and responds to both visual and vestibular motion cues. B) The hippocampus and parahippocampul regions are involved with spatial orientation and navigation functions. All receive vestibular signals regarding body and head motion.

and then used for spatial orientation and navigation, such as driving to the store, or walking through a dark house (McNaughton, Battaglia, Jensen, Moser, & Moser, 2006). Cells in the limbic system and the hippocampus that contribute to these functions have been identified, including place cells, grid cells, and head direction cells (Figure 6B). Place cells in the hippocampus encode specific locations in the environment (O'Keefe, 1976). Grid cells in the entorhinal cortex encode spatial maps in a tessellated pattern (Hafting, Fyhn, Molden, Moser, & Moser, 2005). Head direction cells in the anterior-dorsal thalamus encode heading direction, independent of spatial location (Taube, 1995). It is thought that these cell types work together to provide for spatial orientation, spatial memory, and our ability to navigate. Both place cells and head direction cells depend upon a functioning vestibular system to maintain their directional and orientation information (Stackman, Clark, & Taube, 2002). The pathway by which vestibular signals reach the navigation network is not well understood; however, damage to the vestibular system, hippocampus, and dorsal thalamus regions often disrupts our ability to orient in familiar environments, navigate from place to place, or even to find our way home.

Motion sickness

A common treatment of motion sickness is Dramamine, which helps to reduce the sensitivity of input from your vestibular system to the rest of your body. [Image: Mike Baird, https://goo.gl/zfeqqr, CC BY 2.0, https://goo.gl/BRvSA7]

Although a number of conditions can produce motion sickness, it is generally thought that it is evoked from a mismatch in sensory cues between vestibular, visual, and proprioceptive signals (Yates, Miller, & Lucot, 1998). For example, reading a book in a car on a winding road can produce motion sickness, whereby the accelerations experienced by the vestibular system do not match the visual input. However, if one looks out the window at the scenery going by during the same travel, no sickness occurs because the visual and vestibular cues are in alignment. Sea sickness, a form of motion sickness, appears to be a special case and arises from unusual vertical oscillatory and roll motion. Human studies have found that low frequency oscillations of 0.2 Hz and large amplitudes (such as found in large seas during a storm) are most likely to cause motion sickness, with higher frequencies offering little problems.

Summary

Here, we have seen that the vestibular system transduces and encodes signals about head motion and position with respect to gravity, information that is then used by the brain for many essential functions and behaviors. We actually understand a great deal regarding vestibular contributions to fundamental reflexes, such as compensatory eye movements and balance during motion. More recent progress has been made toward understanding how vestibular signals combine with other sensory cues, such as vision, in the thalamus and cortex to give rise to motion perception. However, there are many complex cognitive abilities that we know require vestibular information to function, such as spatial orientation and navigation behaviors, but these systems are only just beginning to be investigated. Future research regarding vestibular system function will likely be geared to seeking answers to questions regarding how the brain copes with vestibular signal loss. In fact, according to the National Institutes of Health, nearly 35% of Americans over the age of 40 (69 million people) have reported chronic vestibular-related problems. It is therefore of significant importance to human health to better understand how vestibular cues contribute to common brain functions and how better treatment options for vestibular dysfunction can be realized.

Outside Resources

Animated Video of the Vestibular System
http://sites.sinauer.com/neuroscience5e/animations14.01.html

Discussion Questions

1. If a person sustains loss of the vestibular receptors in one ear due to disease or trauma, what symptoms would the person suffer? Would the symptoms be permanent?
2. Often motion sickness is relieved when a person looks at far distance objects, such as things located on the far horizon. Why does far distance viewing help in motion sickness while close distance view (like reading a map or book) make it worse?
3. Vestibular signals combine with visual signals in certain areas of cortex and assist in motion perception. What types of cues does the visual system provide for self motion through space? What types of vestibular signals would be consistent with rotational versus translational motion?

Vocabulary

Abducens nucleus
A group of excitatory motor neurons in the medial brainstem that send projections through the VIth cranial nerve to control the ipsilateral lateral rectus muscle. In addition, abducens interneurons send an excitatory projection across the midline to a subdivision of cells in the ipsilateral oculomotor nucleus, which project through the IIIrd cranial nerve to innervate the ipsilateral medial rectus muscle.

Acetylcholine
An organic compound neurotransmitter consisting of acetic acid and choline. Depending upon the receptor type, acetycholine can have excitatory, inhibitory, or modulatory effects.

Afferent nerve fibers
Single neurons that innervate the receptor hair cells and carry vestibular signals to the brain as part of the vestibulocochlear nerve (cranial nerve VIII).

Aspartate
An excitatory amino acid neurotransmitter that is widely used by vestibular receptors, afferents, and many neurons in the brain.

Compensatory reflexes
A stabilizing motor reflex that occurs in response to a perceived movement, such as the vestibuloocular reflex, or the postural responses that occur during running or skiing.

Depolarized
When receptor hair cells have mechanically gated channels open, the cell increases its membrane voltage, which produces a release of neurotransmitter to excite the innervating nerve fiber.

Detection thresholds
The smallest amount of head motion that can be reliably reported by an observer.

Directional tuning
The preferred direction of motion that hair cells and afferents exhibit where a peak excitatory response occurs and the least preferred direction where no response occurs. Cells are said to be “tuned” for a best and worst direction of motion, with in-between motion directions eliciting a lesser but observable response.

Gamma-aminobutyric acid
A major inhibitory neurotransmitter in the vestibular commissural system.

Gaze stability
A combination of eye, neck, and head responses that are all coordinated to maintain visual fixation (fovea) upon a point of interest.

Glutamate
An excitatory amino acid neurotransmitter that is widely used by vestibular receptors, afferents, and many neurons in the brain.

Hair cells
The receptor cells of the vestibular system. They are termed hair cells due to the many hairlike cilia that extend from the apical surface of the cell into the gelatin membrane. Mechanical gated ion channels in the tips of the cilia open and close as the cilia bend to cause membrane voltage changes in the hair cell that are proportional to the intensity and direction of motion.

Hyperpolarizes
When receptor hair cells have mechanically gated channels close, the cell decreases its membrane voltage, which produces less release of neurotransmitters to inhibit the innervating nerve fiber.

Lateral rectus muscle
An eye muscle that turns outward in the horizontal plane.

Lateral vestibulo-spinal tract
Vestibular neurons that project to all levels of the spinal cord on the ipsilateral side to control posture and balance movements.

Mechanically gated ion channels
Ion channels located in the tips of the stereocilia on the receptor cells that open/close as the cilia bend toward the tallest/smallest cilia, respectively. These channels are permeable to potassium ions, which are abundant in the fluid bathing the top of the hair cells.

Medial vestibulo-spinal tract
Vestibular nucleus neurons project bilaterally to cervical spinal motor neurons for head and neck movement control. The tract principally functions in gaze direction and stability during motion.

Neurotransmitters

A chemical compound used to send signals from a receptor cell to a neuron, or from one neuron to another. Neurotransmitters can be excitatory, inhibitory, or modulatory and are packaged in small vesicles that are released from the end terminals of cells.

Oculomotor nuclei

Includes three neuronal groups in the brainstem, the abducens nucleus, the oculomotor nucleus, and the trochlear nucleus, whose cells send motor commands to the six pairs of eye muscles.

Oculomotor nucleus

A group of cells in the middle brainstem that contain subgroups of neurons that project to the medial rectus, inferior oblique, inferior rectus, and superior rectus muscles of the eyes through the 3rd cranial nerve.

Otoconia

Small calcium carbonate particles that are packed in a layer on top of the gelatin membrane that covers the otolith receptor hair cell stereocilia.

Otolith receptors

Two inner ear vestibular receptors (utricle and saccule) that transduce linear accelerations and head tilt relative to gravity into neural signals that are then transferred to the brain.

Proprioceptive

Sensory information regarding muscle position and movement arising from receptors in the muscles, tendons, and joints.

Semicircular canals

A set of three inner ear vestibular receptors (horizontal, anterior, posterior) that transduce head rotational accelerations into head rotational velocity signals that are then transferred to the brain. There are three semicircular canals in each ear, with the major planes of each canal being orthogonal to each other.

Stereocilia

Hairlike projections from the top of the receptor hair cells. The stereocilia are arranged in ascending height and when displaced toward the tallest cilia, the mechanical gated channels open and the cell is excited (depolarized). When the stereocilia are displaced toward the smallest cilia, the channels close and the cell is inhibited (hyperpolarized).

Torsion
A rotational eye movement around the line of sight that consists of a clockwise or counterclockwise direction.

Vergence angle
The angle between the line of sight for the two eyes. Low vergence angles indicate far-viewing objects, whereas large angles indicate viewing of near objects.

Vestibular compensation
Following injury to one side of vestibular receptors or the vestibulocochlear nerve, the central vestibular nuclei neurons gradually recover much of their function through plasticity mechanisms. The recovery is never complete, however, and extreme motion environments can lead to dizziness, nausea, problems with balance, and spatial memory.

Vestibular efferents
Nerve fibers originating from a nucleus in the brainstem that project from the brain to innervate the vestibular receptor hair cells and afferent nerve terminals. Efferents have a modulatory role on their targets, which is not well understood.

Vestibular system
Consists of a set of motion and gravity detection receptors in the inner ear, a set of primary nuclei in the brainstem, and a network of pathways carrying motion and gravity signals to many regions of the brain.

Vestibulocochlear nerve
The VIIIth cranial nerve that carries fibers innervating the vestibular receptors and the cochlea.

Vestibuloocular reflex
Eye movements produced by the vestibular brainstem that are equal in magnitude and opposite in direction to head motion. The VOR functions to maintain visual stability on a point of interest and is nearly perfect for all natural head movements.

References

Angelaki, D. E., McHenry, M. Q., & Hess, B. J. (2000). Primate translational vestibuloocular reflexes. I. High-frequency dynamics and three-dimensional properties during lateral motion. *J. Neurophysiol 83*, 1637–1647.

Angelaki, D. E., Shaikh, A. G., Green, A. M., & Dickman, J. D. (2004). Neurons compute internal models of the physical laws of motion. *Nature, 430*, 560–564.

Anniko, M., & Arnold, W. (1991). Acetylcholine receptor localization in human adult cochlear and vestibular hair cells. *AC and Otolar, 111*, 491–499.

Beraneck, M., Hachemaoui, M., Idoux, E., Ris, L., Uno, A., Godaux, E., . . . Vibert, N. (2003). Long-term plasticity of ipsilesional medial vestibular nucleus neurons after unilateral labyrinthectomy. *J Neurophysiol, 90*, 184–203.

Blanks, R.H.I., Curthoys, I. S., Bennett, M. L., & Markham, C. H. (1985). Planar relationships of the semicircular canals in rhesus and squirrel monkeys. *Brain Res, 340*, 315–324.

Boyle, R., Carey, J. P., & Highstein, S.M. (1991). Morphological correlates of response dynamics and efferent stimulation in horizontal semicircular canal afferents of the toadfish, *Opsanus tau. J Neurophysiol, 66*, 1504–1511.

Brodal, A. The vestibular nuclei in the macaque monkey. (1984). *J. Comp. Neurol., 227*, 252–266.

Carleton, S. C., & Carpenter, M. B. (1984). Distribution of primary vestibular fibers in the brainstem and cerebellum of the monkey. *Brain Res, 294*, 281–298.

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011). A comparison of vestibular spatiotemporal tuning in macaque parietoinsular vestibular cortex, ventral intraparietal area, and medial superior temporal area. *J. Neurosci., 31*, 3082–3094.

Crane, B. T., & Demer, J. L. (1997). Human gaze stabilization during natural activities: translation, rotation, magnification, and target distance effects. *J Neurophysiol, 78*, 2129–2144.

Dickman, J. D., & Angelaki, D. E. (2002). Vestibular convergence patterns in vestibular nuclei neurons of alert primates. *J. Neurophysiol., 88*, 3518–3533.

Dickman, J. D., & Correia, M. J. (1989). Responses of pigeon horizontal semicircular canal afferent fibers. I. Step, trapezoid, and low frequency sinusoid mechanical and rotational stimulation. *J Neurophysiol, 62*, 1090–1101.

Dickman, J. D., & Fang, Q. (1996). Differential central projections of vestibular afferents in pigeons. *J Comp Neurol, 367*, 110–121.

Einstein, A. (1907).On the relativity principle and the conclusions drawn from it. *Jahrb. Radioakt, 4*, 411–462.

Ezure, K., & Graf, K.W. (1984). A quantitative analysis of the spatial organization of the vestibulo-ocular reflexes in lateral- and frontal-eyed animals. I. Orientation of semicircular canals and extraocular muscles. *Neuroscience, 12*, 85–94.

Fernandez, C., & Goldberg, J. M. (1976b). Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. II. Directional selectivity and force-response relations. *J. Neurophysiol., 39*, 985–995.

Fernandez, C., & Goldberg, J. M. (1971). Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey. II. Response to sinusoidal stimulation and dynamics of peripheral vestibular system. *J Neurophysiol, 34*, 661–675.

Fernandez, C., & Goldberg, J.M. (1976a). Physiology of peripheral neurons innervating otolith organ in the squirrel monkey. I. Response to static tilts and to long-duration centrifugal force. *J. Neurophysiol., 39*, 970–984.

Fetter, M., & Zee, D. S. (1988). Recovery from unilateral labyrinthectomy in rhesus monkey. *J Neurophysiol, 59*, 370–393.

Flock, A. (1964). Structure of the macula utriculi with special reference to directional interplay of sensory responses as revealed by morphological polarization. *J Cell Biol, 22*, 413–431.

Fukushima, K., Sato, T., Fukushima, J. Shinmei, Y., and Kaneko C.R. (2000). Activity of smooth pursuit related neurons in the monkey perarcuate cortes during pursuit and passive whole-body rotation. *J. Neurophys., 83*, 563–587.

Grusser, O., Pause, M., & Schreiter, U. (1982). Neuronal responses in the parieto-insular vestibular cortex of alert java monkeys (Macaca fascicularis). In A. Roucoux and M. Crommelinck (Eds.), *Physiological and pathological aspects of eye movements*. The Haque: W. Junk Publishers.

Gu, Y., DeAngelis, G. C., & Angelaki, D. E. (2007). A functional link between area mstd and heading perception based on vestibular signals. *Nat Neurosci, 10*, 1038–1047.

Gu, Y., Watkins, P. V., Angelaki, D. E., & DeAngelis, G. C. (2006). Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci, 26*, 73–85.

Guldin, W. O., Akbarian, S., & Grusser, O. J. (1992). Cortico-cortical connections and cytoarchitectonics of the primate vestibular cortex: A study in squirrel monkeys (Saimiri sciureus). *J. Comp. Neurol*., 326, 375–401.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature, 436*, 801–806.

Iurato, S. (1967). *Submicroscopic structure of the inner ear*. London, England: Pergamon Press.

Kasahara, M., & Uchino, Y. (1974). Bilateral semicircular canal inputs to neurons in cat vestibular nuclei. *Exp Brain Res, 20*, 285–296.

Keshner, E. A., & Peterson, B. W. (1995). Mechanisms controlling human head stabilization. I. Head-neck dynamics during random rotations in the horizontal plane. *J Neurophysiol, 73*, 2293–2301.

Lindeman, H. H. (1969). Studies on the morphology of the sensory regions of the vestibular apparatus with 45 figures. *Ergeb Anat Entwicklungsgesch, 42*, 1–113.

Lisberger, S. G., Pavelko, T. A., & Broussard, D. M. (1994). Neural basis for motor learning in the vestibuloocular reflex of primates. I. Changes in the responses of brain stem neurons. *J. Neurophysiol, 72*, 928–973.

Marlinski, V., & McCrea, R. A. (2008). Activity of ventroposterior thalamus neurons during rotation and translation in the horizontal plane in the alert squirrel monkey. *J Neurophysiol, 99*, 2533–2545.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-.B. (2006). Path integration and the neural basis of the "cognitive map." *Nature Rev., 7*, 663–678.

Meng, H., May, P. J., Dickman, J. D., & Angelaki, D. E. (2007). Vestibular signals in primate thalamus: Properties and origins. *J Neurosci, 27*, 13590–13602.

Newlands, S. D., & Perachio, A. A. (1990). Compensation of horizontal canal related activity in the medial vestibular nucleus following unilateral labyrinth ablation in the decerebrate gerbil. i. type i neurons. *Exp Brain Res, 82*, 359–372.

Newlands, S. D., Hesse, S. V., Haque, A., & Angelaki, D. E. (2001). Head unrestrained horizontal gaze shifts after unilateral labyrinthectomy in the rhesus monkey. *Exp Brain Res, 140*, 25–33.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Exp Neurol, 51*, 78–109.

Paige, G. D., & Tomko, D. L. (1991). Eye movement responses to linear head motion in the squirrel monkey. I. Basic characteristics. *J Neurophysiol, 65*, 1170–1182.

Peterson, B. W., Goldberg, J., Bilotto, G., & Fuller, J. H. (1985). Cervicocollic reflex: Its dynamic properties and interaction with vestibular reflexes. *J Neurophysiol, 54*, 90–109.

Precht, W., & Shimazu, H. (1965). Functional connections of tonic and kinetic vestibular neurons with primary vestibular afferents. *J Neurophysiol, 28*, 1014–1028.

Sadeghi, S. G., Chacron, M. J., Taylor, M. C., & Cullen, K. E. (2007). Neural variability, detection thresholds, and information transmission in the vestibular system. *J. Neurosci., 27*, 771–781.

Schwarz, U., Busettini, C., & Miles, F. A. (1989). Ocular responses to linear motion are inversely proportional to viewing distance. *Science, 245*, 1394–1396.

Shimazu, H., & Precht, W. (1966). Inhibition of central vestibular neurons from the contralateral labyrinth and its mediating pathway. *J Neurophysiol, 29*, 467–492.

Shinoda, Y., Sugiuchi, Y., Futami, T., Ando, N., & Kawasaki, T. (1994). Input patterns and pathways from the six semicircular canals to motoneurons of neck muscles. I. The multifidus muscle group. *J. Neurophysiol., 72*, 2691–2802.

Shotwell, S. L, Jacobs, R., & Hudspeth, A. J. (1981). Directional sensitivity of individual vertebrate hair cells to controlled deflection of their hair bundles. *ANYAS, 374*, 1–10.

Stackman, R. W., Clark, A. S., & Taube, J. (2002). Hippocampal spatial representations require vestibular input. *Hippocampus, 12*, 291–303.

Taube, J .S. (1995). Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J. Neurosci., 15*, 70–86.

Van Egmond, A. A. J., Groen, J. J., & Jongkess, L. B. W. (1949). The mechanics of the semicircular canals. *J Physiol Lond, 110*, 1–17.

Wersäll, J. (1956). Studies on the structure and innervation of the sensory epithelium of the cristae ampullares in the guinea pig; a light and electron microscopic investigation. *Acta Otolaryngol Supp, 126*, 1–85.

Wilson, V. J., Boyle, R., Fukushima, K., Rose, P. K., Shinoda, Y., Sugiuchi, Y., & Uchino, Y. (1995). The vestibulocollic reflex. *J Vestib Res, 5*, 147–170.

Yates, B. J., Miller A. D., & Lucot, J. B. (1998). Physiological basis and pharmacology of motion sickness: An update. *Brain Res Bull, 47*, 395–406.

Yu, X. J., Dickman, J. D., Angelaki, D. E. (2012). Detection thresholds of macaque otolith afferents. *J Neurosci., 32*, 8306–8316.

Zhou, W., Tang, B.F., Newlands, S.D., King, W.M. (2006). Responses of monkey vestibular-only neurons to translation and angular rotation. *J Neurophysiol 96*, 2915-2930.

17
Multi-Modal Perception

Lorin Lachs

Most of the time, we perceive the world as a unified bundle of sensations from multiple sensory modalities. In other words, our perception is multimodal. This module provides an overview of multimodal perception, including information about its neurobiology and its psychological effects.

Learning Objectives

- Define the basic terminology and basic principles of multimodal perception.
- Describe the neuroanatomy of multisensory integration and name some of the regions of the cortex and midbrain that have been implicated in multisensory processing.
- Explain the difference between multimodal phenomena and crossmodal phenomena.
- Give examples of multimodal and crossmodal behavioral effects.

Perception: Unified

Although it has been traditional to study the various senses independently, most of the time, perception operates in the context of information supplied by multiple **sensory modalities** at the same time. For example, imagine if you witnessed a car collision. You could describe the stimulus generated by this event by considering each of the senses independently; that is, as a set of **unimodal** stimuli. Your eyes would be stimulated with patterns of light energy bouncing off the cars involved. Your ears would be stimulated with patterns of acoustic energy emanating from the collision. Your nose might even be stimulated by the smell of burning rubber or gasoline.

If you were a witness to this scene you'd be able to describe it using input from many of your senses. Your experience would be multimodal. [Image: Photo Grrrrr, https://goo.gl/dzfKs8, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

However, all of this information would be relevant to the same thing: your perception of the car collision. Indeed, unless someone was to explicitly ask you to describe your perception in unimodal terms, you would most likely experience the event as a unified bundle of sensations from multiple senses. In other words, your perception would be **multimodal**. The question is whether the various sources of information involved in this multimodal stimulus are processed separately by the perceptual system or not.

For the last few decades, perceptual research has pointed to the importance of **multimodal perception**: the effects on the perception of events and objects in the world that are observed when there is information from more than one sensory modality. Most of this research indicates that, at some point in perceptual processing, information from the various sensory modalities is **integrated**. In other words, the information is combined and treated as a unitary representation of the world.

Questions About Multimodal Perception

Several theoretical problems are raised by multimodal perception. After all, the world is a “blooming, buzzing world of confusion” that constantly bombards our perceptual system with light, sound, heat, pressure, and so forth. To make matters more complicated, these stimuli come from multiple events spread out over both space and time. To return to our example: Let’s say the car crash you observed happened on Main Street in your town. Your perception during the car crash might include a lot of stimulation that was *not* relevant to the car crash. For example, you might also overhear the conversation of a nearby couple, see a bird flying into a tree, or smell the delicious scent of freshly baked bread from a nearby bakery (or all three!). However, you would most likely not make the mistake of associating any of these stimuli with the car crash. In fact, we rarely combine the auditory stimuli associated with one event with the visual stimuli associated with another (although, under some unique circumstances—such as ventriloquism—we do). How is the brain able to take the information from separate sensory modalities and match it appropriately, so that stimuli that belong together stay together, while stimuli that do not belong together get treated separately? In other words, how does the perceptual system determine which unimodal stimuli must be integrated, and which must not?

Once unimodal stimuli have been appropriately integrated, we can further ask about the consequences of this integration: What are the effects of multimodal perception that would not be present if perceptual processing were only unimodal? Perhaps the most robust finding in the study of multimodal perception concerns this last question. No matter whether you are looking at the actions of neurons or the behavior of individuals, it has been found that responses to multimodal stimuli are typically greater than the combined response to either modality independently. In other words, if you presented the stimulus in one modality at a time and measured the response to each of these unimodal stimuli, you would find that adding them together would still not equal the response to the multimodal stimulus. This **superadditive effect of multisensory integration** indicates that there are consequences resulting from the integrated processing of multimodal stimuli.

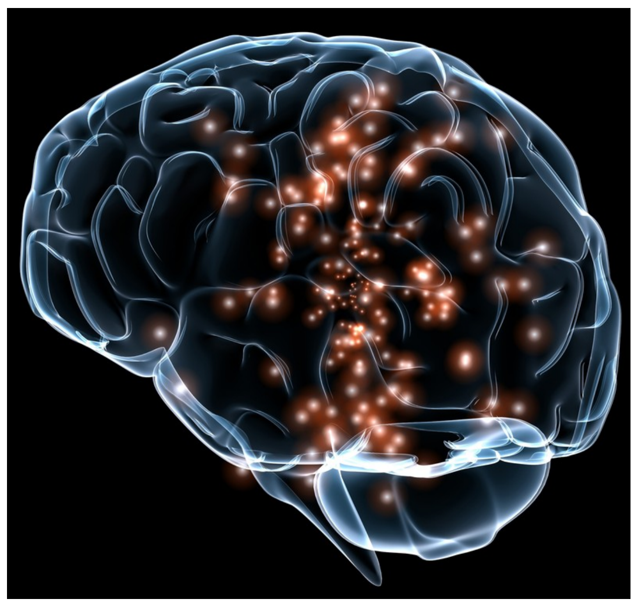
The extent of the superadditive effect (sometimes referred to as **multisensory enhancement**) is determined by the strength of the response to the single stimulus modality with the biggest effect. To understand this concept, imagine someone speaking to you in a noisy environment (such as a crowded party). When discussing this type of multimodal stimulus, it is often useful to describe it in terms of its **unimodal components**: In this case, there is an auditory component (the sounds generated by the speech of the person speaking to you) and a visual component (the visual form of the face movements as the person speaks to you). In the crowded party, the auditory component of the person's speech might be difficult to process (because of the surrounding party noise). The potential for visual information about speech—lipreading—to help in understanding the speaker's message is, in this situation, quite large. However, if you were listening to that same person speak in a quiet library, the auditory portion would probably be sufficient for receiving the message, and the visual portion would help very little, if at all (Sumby & Pollack, 1954). In general, for a stimulus with multimodal components, if the response to each component (on its own) is weak, then the opportunity for multisensory enhancement is very large. However, if one component—by itself—is sufficient to evoke a strong response, then the opportunity for multisensory enhancement is relatively small. This finding is called the **Principle of Inverse Effectiveness** (Stein & Meredith, 1993) because the effectiveness of multisensory enhancement is inversely related to the unimodal response with the greatest effect.

Another important theoretical question about multimodal perception concerns the neurobiology that supports it. After all, at some point, the information from each sensory modality is definitely separated (e.g., light comes in through the eyes, and sound comes in through the ears). How does the brain take information from different neural systems (optic, auditory, etc.) and combine it? If our experience of the world is multimodal, then it must be the case that at some point during perceptual processing, the unimodal information coming from separate sensory organs—such as the eyes, ears, skin—is combined. A related question asks where in the brain this integration takes place. We turn to these questions in the next section.

Biological Bases of Multimodal Perception

Multisensory Neurons and Neural Convergence

A surprisingly large number of brain regions in the midbrain and cerebral cortex are related to multimodal perception. These regions contain neurons that respond to stimuli from not just one, but multiple sensory modalities. For example, a region called the superior temporal sulcus contains single neurons that respond to both the visual and auditory components of speech (Calvert, 2001; Calvert, Hansen, Iversen, & Brammer, 2001). These multisensory convergence zones are interesting, because they are a kind of neural intersection of information coming from the different senses. That is, neurons that are devoted to the processing of one sense at a time—say vision or touch—send their information to the convergence zones, where it is processed together.



In order for us to perceive the world effectively, neurons from our various senses carry information that is integrated in the brain. [Image: DARPA, https://goo.gl/kat7ws, CC0 Public Domain, https://goo.gl/m25gce]

One of the most closely studied multisensory convergence zones is the superior colliculus (Stein & Meredith, 1993), which receives inputs from many different areas of the brain, including regions involved in the unimodal processing of visual and auditory stimuli (Edwards, Ginsburgh, Henkel, & Stein, 1979). Interestingly, the superior colliculus is involved in the "orienting response," which is the behavior associated with moving one's eye gaze toward the location of a seen or heard stimulus. Given this function for the superior colliculus, it is hardly surprising that there are multisensory neurons found there (Stein & Stanford, 2008).

Crossmodal Receptive Fields

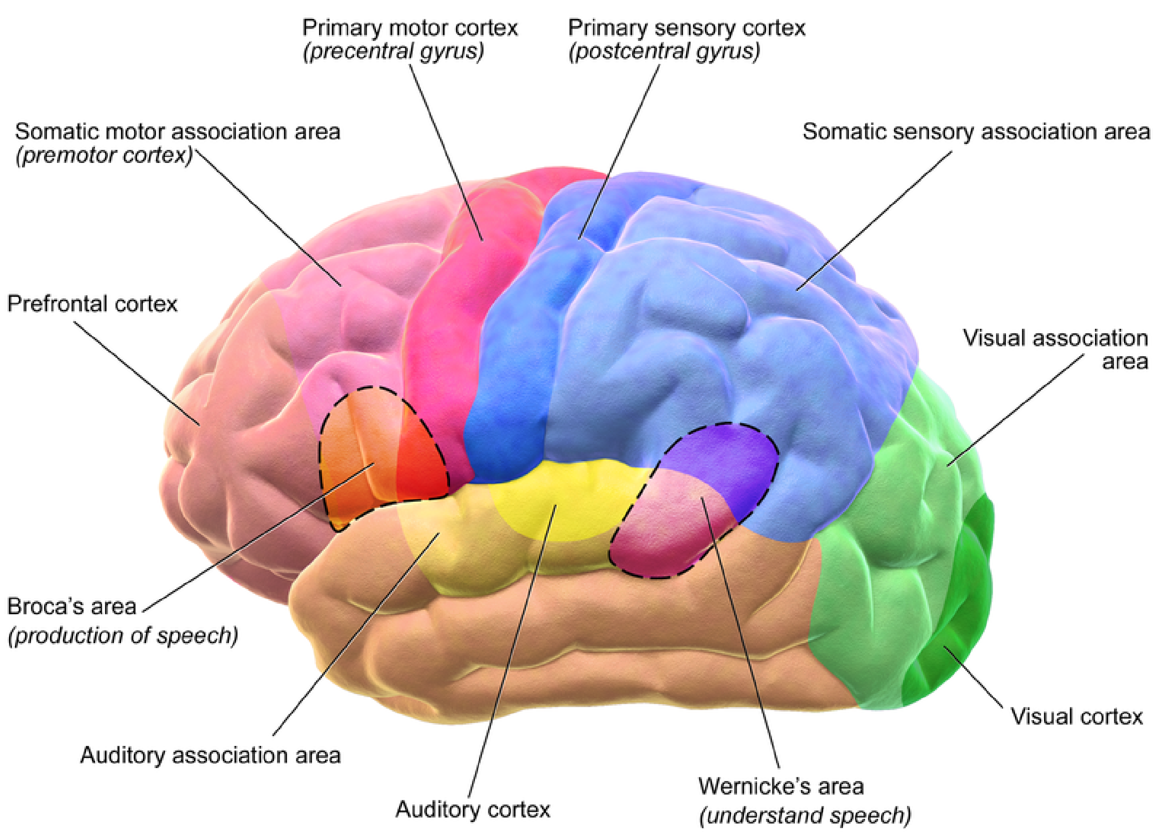
The details of the anatomy and function of multisensory neurons help to answer the question of how the brain integrates stimuli appropriately. In order to understand the details, we need to discuss a neuron's receptive field. All over the brain, neurons can be found that respond only to stimuli presented in a very specific region of the space immediately surrounding the perceiver. That region is called the neuron's **receptive field**. If a stimulus is presented in a neuron's receptive field, then that neuron responds by increasing or decreasing its firing rate. If a stimulus is presented outside of a neuron's receptive field, then there is no effect on the neuron's firing rate. Importantly, when two neurons send their information to a third neuron, the third neuron's receptive field is the combination of the receptive fields of the two input neurons. This is called neural convergence, because the information from multiple neurons converges on a single neuron. In the case of multisensory neurons, the convergence arrives from different sensory modalities. Thus, the receptive fields of multisensory neurons are the combination of the receptive fields of neurons located in different sensory pathways.

Now, it could be the case that the neural convergence that results in multisensory neurons is set up in a way that ignores the locations of the input neurons' receptive fields. Amazingly, however, these **crossmodal receptive fields** overlap. For example, a multisensory neuron in the superior colliculus might receive input from two unimodal neurons: one with a visual receptive field and one with an auditory receptive field. It has been found that the unimodal receptive fields refer to the same locations in space—that is, the two unimodal neurons respond to stimuli in the same region of space. Crucially, the overlap in the crossmodal receptive fields plays a vital role in the integration of **crossmodal stimuli**. When the information from the separate modalities is coming from within these overlapping receptive fields, then it is treated as having come from the same location—and the neuron responds with a superadditive (enhanced) response. So, part of the information that is used by the brain to combine multimodal inputs is the location in space from which the stimuli came.

This pattern is common across many multisensory neurons in multiple regions of the brain. Because of this, researchers have defined the **spatial principle of multisensory integration**: Multisensory enhancement is observed when the sources of stimulation are spatially related to one another. A related phenomenon concerns the *timing* of crossmodal stimuli. Enhancement effects are observed in multisensory neurons only when the inputs from different senses arrive within a short time of one another (e.g., Recanzone, 2003).

Multimodal Processing in Unimodal Cortex

Multisensory neurons have also been observed outside of multisensory convergence zones, in areas of the brain that were once thought to be dedicated to the processing of a single modality (unimodal cortex). For example, the primary visual cortex was long thought to be devoted to the processing of exclusively visual information. The primary visual cortex is the first stop in the cortex for information arriving from the eyes, so it processes very low-level information like edges. Interestingly, neurons have been found in the primary visual cortex that receives information from the primary auditory cortex (where sound information from the auditory pathway is processed) and from the superior temporal sulcus (a multisensory convergence zone mentioned above). This is remarkable because it indicates that the processing of visual information is, from a very early stage, influenced by auditory information.



There are zones in the human brain where sensory information comes together and is integrated such as the Auditory, Visual and Motor Cortices pictured here. [Image: BruceBlaus, https://goo.gl/UqKBI3, CC BY 3.0, https://goo.gl/b58TcB]

There may be two ways for these multimodal interactions to occur. First, it could be that the processing of auditory information in relatively late stages of processing feeds back to influence low-level processing of visual information in unimodal cortex (McDonald, Teder-Sälejärvi, Russo, & Hillyard, 2003). Alternatively, it may be that areas of unimodal cortex contact each other directly (Driver & Noesselt, 2008; Macaluso & Driver, 2005), such that multimodal integration is a fundamental component of all sensory processing.

In fact, the large numbers of multisensory neurons distributed all around the cortex—in multisensory convergence areas and in primary cortices—has led some researchers to propose that a drastic reconceptualization of the brain is necessary (Ghazanfar & Schroeder, 2006). They argue that the cortex should not be considered as being divided into isolated regions that process only one kind of sensory information. Rather, they propose that these areas only *prefer* to process information from specific modalities but engage in low-level multisensory processing whenever it is beneficial to the perceiver (Vasconcelos et al., 2011).

Behavioral Effects of Multimodal Perception

Although neuroscientists tend to study very simple interactions between neurons, the fact that they've found so many crossmodal areas of the cortex seems to hint that the way we experience the world is fundamentally multimodal. As discussed above, our intuitions about perception are consistent with this; it does not seem as though our perception of events is constrained to the perception of each sensory modality independently. Rather, we perceive a unified world, regardless of the sensory modality through which we perceive it.

It will probably require many more years of research before neuroscientists uncover all the details of the neural machinery involved in this unified experience. In the meantime, experimental psychologists have contributed to our understanding of multimodal perception through investigations of the behavioral effects associated with it. These effects fall into two broad classes. The first class—**multimodal phenomena**—concerns the binding of inputs from multiple sensory modalities and the effects of this binding on perception. The second class—**crossmodal phenomena**—concerns the influence of one sensory modality on the perception of another (Spence, Senkowski, & Roder, 2009).

Multimodal Phenomena

Audiovisual Speech

Multimodal phenomena concern stimuli that generate simultaneous (or nearly simultaneous) information in more than one sensory modality. As discussed above, speech is a classic example of this kind of stimulus. When an individual speaks, she generates sound waves that carry meaningful information. If the perceiver is also looking at the speaker, then that perceiver also has access to *visual* patterns that carry meaningful information. Of course, as anyone who has ever tried to lipread knows, there are limits on how informative visual speech information is. Even so, the visual speech pattern alone is sufficient for very robust speech perception. Most people assume that deaf individuals are much better at lipreading than individuals with normal hearing. It may come as a surprise to learn, however, that some individuals with normal hearing are also remarkably good at lipreading (sometimes called "speechreading"). In fact, there is a wide range of speechreading ability in both normal hearing and deaf populations (Andersson, Lyxell, Rönnberg, & Spens, 2001). However, the reasons for this wide range of performance are not well understood (Auer & Bernstein, 2007; Bernstein, 2006; Bernstein, Auer, & Tucker, 2001; Mohammed et al., 2005).

In a noisy and poorly lit environment such as a nightclub in order to have a conversation we rely on audiovisual speech to understand others. [Image: Jeremy Keith, https://goo.gl/18sLfg, CC BY 2.0, https://goo.gl/v4Y0Zv]

How does visual information about speech interact with auditory information about speech? One of the earliest investigations of this question examined the accuracy of recognizing spoken words presented in a noisy context, much like in the example above about talking at a crowded party. To study this phenomenon experimentally, some irrelevant noise ("white noise"—which sounds like a radio tuned between stations) was presented to participants. Embedded in the white noise were spoken words, and the participants' task was to identify the words. There were two conditions: one in which only the auditory component of the words was presented (the "auditory-alone" condition), and one in both the auditory and visual components were presented (the "audiovisual" condition). The noise levels were also varied, so that on some trials, the noise was very loud relative to the loudness of the words, and on other trials, the noise was very soft relative to the words. Sumby and Pollack (1954) found that the accuracy of identifying the spoken words was much higher for the audiovisual condition than it was in the auditory-alone condition. In addition, the pattern of results was consistent with the Principle of Inverse Effectiveness: The advantage gained by audiovisual presentation was highest when the auditory-alone condition performance was lowest (i.e., when the noise was loudest). At these noise levels, the audiovisual advantage was considerable: It was estimated that allowing the participant to see the speaker was equivalent to turning the volume of the noise down by over half. Clearly, the audiovisual advantage can have dramatic effects on behavior.

Another phenomenon using audiovisual speech is a very famous illusion called the "**McGurk effect**" (named after one of its discoverers). In the classic formulation of the illusion, a movie is recorded of a speaker saying the syllables "gaga." Another movie is made of the same speaker saying the syllables "baba." Then, the auditory portion of the "baba" movie is dubbed onto the visual portion of the "gaga" movie. This combined stimulus is presented to participants, who are asked to report what the speaker in the movie said. McGurk and MacDonald (1976) reported that 98 percent of their participants reported hearing the syllable "dada"—which was in neither the visual nor the auditory components of the stimulus. These results indicate that when visual and auditory information about speech is integrated, it can have profound effects on perception.

Tactile/Visual Interactions in Body Ownership

Not all multisensory integration phenomena concern speech, however. One particularly compelling multisensory illusion involves the integration of tactile and visual information in the perception of body ownership. In the "**rubber hand illusion**" (Botvinick & Cohen, 1998), an observer is situated so that one of his hands is not visible. A fake rubber hand is placed near the obscured hand, but in a visible location. The experimenter then uses a light paintbrush to simultaneously stroke the obscured hand and the rubber hand in the same locations. For example, if the middle finger of the obscured hand is being brushed, then the middle finger of the rubber hand will also be brushed. This sets up a correspondence between the tactile sensations (coming from the obscured hand) and the visual sensations (of the rubber hand). After a short time (around 10 minutes), participants report feeling as though the rubber hand "belongs" to them; that is, that the rubber hand is a part of their body. This feeling can be so strong that surprising the participant by hitting the rubber hand with a hammer often leads to a reflexive withdrawing of the obscured hand—even though it is in no danger at all. It appears, then, that our awareness of our own bodies may be the result of multisensory integration.

Crossmodal Phenomena

Crossmodal phenomena are distinguished from multimodal phenomena in that they concern the influence one sensory modality has on the perception of another.

Visual Influence on Auditory Localization

Ventriloquists are able to trick us into believing that what we see and what we hear are the same where, in truth, they are not. [Image: Amanda Ferrell, CC0 Public Domain, https://goo.gl/m25gce]

A famous (and commonly experienced) crossmodal illusion is referred to as "the ventriloquism effect." When a ventriloquist appears to make a puppet speak, she fools the listener into thinking that the location of the origin of the speech sounds is at the puppet's mouth. In other words, instead of localizing the auditory signal (coming from the mouth of a ventriloquist) to the correct place, our perceptual system localizes it incorrectly (to the mouth of the puppet).

Why might this happen? Consider the information available to the observer about the location of the two components of the stimulus: the sounds from the ventriloquist's mouth and the visual movement of the puppet's mouth. Whereas it is very obvious where the visual stimulus is coming from (because you can see it), it is much more difficult to pinpoint the location of the sounds. In other words, the very precise visual location of mouth movement apparently overrides the less well-specified location of the auditory information. More generally, it has been found that the location of a wide variety of auditory stimuli can be affected by the simultaneous presentation of a visual stimulus (Vroomen & De Gelder, 2004). In addition, the ventriloquism effect has been demonstrated for objects in motion: The motion of a visual object can influence the perceived direction of motion of a moving sound source (Soto-Faraco, Kingstone, & Spence, 2003).

Auditory Influence on Visual Perception

A related illusion demonstrates the opposite effect: where sounds have an effect on visual perception. In the double flash illusion, a participant is asked to stare at a central point on a computer monitor. On the extreme edge of the participant's vision, a white circle is briefly flashed one time. There is also a simultaneous auditory event: either one beep or two beeps in rapid succession. Remarkably, participants report seeing *two* visual flashes when the flash is accompanied by two beeps; the same stimulus is seen as a single flash in the context of a single beep or no beep (Shams, Kamitani, & Shimojo, 2000). In other words, the number of heard beeps influences the number of seen flashes!

Another illusion involves **the perception of collisions between two circles (called "balls")** moving toward each other and continuing through each other. Such stimuli can be perceived as either two balls moving through each other or as a collision between the two balls that then bounce off each other in opposite directions. Sekuler, Sekuler, and Lau (1997) showed that the presentation of an auditory stimulus at the time of contact between the two balls strongly influenced the perception of a collision event. In this case, the perceived sound influences the interpretation of the ambiguous visual stimulus.

Crossmodal Speech

Experiments have demonstrated that by simply observing a speaker, with no auditory information, we can gather important clues about the actual sound of their voice. [Ken Whytock, https://goo.gl/VQJssP, CC BY-NC 2.0, https://goo.gl/tgFydH]

Several crossmodal phenomena have also been discovered for speech stimuli. These crossmodal speech effects usually show altered perceptual processing of unimodal stimuli (e.g., acoustic patterns) by virtue of prior experience with the alternate unimodal stimulus (e.g., optical patterns). For example, Rosenblum, Miller, and Sanchez (2007) conducted an experiment examining the ability to become familiar with a person's voice. Their first interesting finding was unimodal: Much like what happens when someone repeatedly hears a person speak, perceivers can become familiar with the "visual voice" of a speaker. That is, they can become familiar with the person's speaking style simply by seeing that person speak. Even more astounding was their crossmodal finding: Familiarity with this *visual* information also led to increased recognition of the speaker's *auditory* speech, to which participants had never had exposure.

Similarly, it has been shown that when perceivers see a speaking face, they can identify the (auditory-alone) voice of that speaker, and vice versa (Kamachi, Hill, Lander, & Vatikiotis-Bateson, 2003; Lachs & Pisoni, 2004a, 2004b, 2004c; Rosenblum, Smith, Nichols, Lee, & Hale, 2006). In other words, the visual form of a speaker engaged in the act of speaking appears to contain information about what that speaker should sound like. Perhaps more surprisingly, the auditory form of speech seems to contain information about what the speaker should look like.

Conclusion

In this module, we have reviewed some of the main evidence and findings concerning the role of multimodal perception in our experience of the world. It appears that our nervous system (and the cortex in particular) contains considerable architecture for the processing of information arriving from multiple senses. Given this neurobiological setup, and the diversity of behavioral phenomena associated with multimodal stimuli, it is likely that the investigation of multimodal perception will continue to be a topic of interest in the field of experimental perception for many years to come.

Outside Resources

Article: A review of the neuroanatomy and methods associated with multimodal perception:
http://dx.doi.org/10.1016/j.neubiorev.2011.04.015

Journal: Experimental Brain Research Special issue: Crossmodal processing
http://www.springerlink.com/content/0014-4819/198/2-3

TED Talk: Optical Illusions
http://www.ted.com/talks/beau_lotto_optical_illusions_show_how_we_see

Video: McGurk demo
http://youtu.be/aFPtc8BVdJk

Video: The Rubber Hand Illusion
http://www.youtube.com/watch?v=sxwn1w7MJvk

Web: Double-flash illusion demo
http://www.cns.atr.jp/~kmtn/soundInducedIllusoryFlash2/

Discussion Questions

1. The extensive network of multisensory areas and neurons in the cortex implies that much perceptual processing occurs in the context of multiple inputs. Could the processing of unimodal information ever be useful? Why or why not?
2. Some researchers have argued that the Principle of Inverse Effectiveness (PoIE) results from ceiling effects: Multisensory enhancement cannot take place when one modality is sufficient for processing because in such cases it is not possible for processing to be enhanced (because performance is already at the "ceiling"). On the other hand, other researchers claim that the PoIE stems from the perceptual system's ability to assess the relative value of stimulus cues, and to use the most reliable sources of information to construct a representation of the outside world. What do you think? Could these two possibilities ever be teased apart? What kinds of experiments might one conduct to try to get at this issue?
3. In the late 17th century, a scientist named William Molyneux asked the famous philosopher

John Locke a question relevant to modern studies of multisensory processing. The question was this: Imagine a person who has been blind since birth, and who is able, by virtue of the sense of touch, to identify three dimensional shapes such as spheres or pyramids. Now imagine that this person suddenly receives the ability to see. Would the person, without using the sense of touch, be able to identify those same shapes visually? Can modern research in multimodal perception help answer this question? Why or why not? How do the studies about crossmodal phenomena inform us about the answer to this question?

Vocabulary

Bouncing balls illusion
The tendency to perceive two circles as bouncing off each other if the moment of their contact is accompanied by an auditory stimulus.

Crossmodal phenomena
Effects that concern the influence of the perception of one sensory modality on the perception of another.

Crossmodal receptive field
A receptive field that can be stimulated by a stimulus from more than one sensory modality.

Crossmodal stimulus
A stimulus with components in multiple sensory modalties that interact with each other.

Double flash illusion
The false perception of two visual flashes when a single flash is accompanied by two auditory beeps.

Integrated
The process by which the perceptual system combines information arising from more than one modality.

McGurk effect
An effect in which conflicting visual and auditory components of a speech stimulus result in an illusory percept.

Multimodal
Of or pertaining to multiple sensory modalities.

Multimodal perception
The effects that concurrent stimulation in more than one sensory modality has on the perception of events and objects in the world.

Multimodal phenomena
Effects that concern the binding of inputs from multiple sensory modalities.

Multisensory convergence zones
Regions in the brain that receive input from multiple unimodal areas processing different sensory modalities.

Multisensory enhancement
See “superadditive effect of multisensory integration.”

Primary auditory cortex
A region of the cortex devoted to the processing of simple auditory information.

Primary visual cortex
A region of the cortex devoted to the processing of simple visual information.

Principle of Inverse Effectiveness
The finding that, in general, for a multimodal stimulus, if the response to each unimodal component (on its own) is weak, then the opportunity for multisensory enhancement is very large. However, if one component—by itself—is sufficient to evoke a strong response, then the effect on the response gained by simultaneously processing the other components of the stimulus will be relatively small.

Receptive field
The portion of the world to which a neuron will respond if an appropriate stimulus is present there.

Rubber hand illusion
The false perception of a fake hand as belonging to a perceiver, due to multimodal sensory information.

Sensory modalities
A type of sense; for example, vision or audition.

Spatial principle of multisensory integration
The finding that the superadditive effects of multisensory integration are observed when the sources of stimulation are spatially related to one another.

Superadditive effect of multisensory integration
The finding that responses to multimodal stimuli are typically greater than the sum of the independent responses to each unimodal component if it were presented on its own.

Unimodal
Of or pertaining to a single sensory modality.

Unimodal components
The parts of a stimulus relevant to one sensory modality at a time.

Unimodal cortex
A region of the brain devoted to the processing of information from a single sensory modality.

References

Andersson, U., Lyxell, B., Rönnberg, J., & Spens, K.-E. (2001). Cognitive correlates of visual speech understanding in hearing-impaired individuals. Journal of Deaf Studies and Deaf Education, 6(2), 103–116. doi: 10.1093/deafed/6.2.103

Auer, E. T., Jr., & Bernstein, L. E. (2007). Enhanced visual speech perception in individuals with early-onset hearing impairment. Journal of Speech, Language, and Hearing Research, 50 (5), 1157–1165. doi: 10.1044/1092-4388(2007/080)

Bernstein, L. E. (2006). Visual speech perception. In E. Vatikiotis-Bateson, G. Bailley, & P. Perrier (Eds.), Audio-visual speech processing. Cambridge, MA: MIT Press.

Bernstein, L. E., Auer, E. T., Jr., & Tucker, P. E. (2001). Enhanced speechreading in deaf adults: Can short-term training/practice close the gap for hearing adults? Journal of Speech, Language, and Hearing Research, 44, 5–18.

Botvinick, M., & Cohen, J. (1998). Rubber hands /`feel/' touch that eyes see. [10.1038/35784]. Nature, 391(6669), 756–756.

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. Cerebral Cortex, 11, 1110–1123.

Calvert, G. A., Hansen, P. C., Iversen, S. D., & Brammer, M. J. (2001). Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the bold effect. NeuroImage, 14(2), 427–438. doi: 10.1006/nimg.2001.0812

Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. Neuron, 57(1), 11–23. doi: 10.1016/j.neuron.2007.12.013

Edwards, S. B., Ginsburgh, C. L., Henkel, C. K., & Stein, B. E. (1979). Sources of subcortical projections to the superior colliculus in the cat. Journal of Comparative Neurology, 184(2), 309–329. doi: 10.1002/cne.901840207

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? TRENDS in Cognitive Sciences, 10(6), 278-285. doi: 10.1016/j.tics.2006.04.008

Kamachi, M., Hill, H., Lander, K., & Vatikiotis-Bateson, E. (2003). "Putting the face to the voice": Matching identity across modality. Current Biology, 13, 1709–1714.

Lachs, L., & Pisoni, D. B. (2004a). Crossmodal source identification in speech perception. Ecological Psychology, 16(3), 159–187.

Lachs, L., & Pisoni, D. B. (2004b). Crossmodal source information and spoken word recognition. Journal of Experimental Psychology: Human Perception & Performance, 30(2), 378–396.

Lachs, L., & Pisoni, D. B. (2004c). Specification of crossmodal source information in isolated kinematic displays of speech. Journal of the Acoustical Society of America, 116(1), 507–518.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: A window onto functional integration in the human brain. Trends in Neurosciences, 28(5), 264–271. doi: 10.1016/j.tins.2005.03.008

McDonald, J. J., Teder-Sälejärvi, W. A., Russo, F. D., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. Journal of Cognitive Neuroscience, 15(1), 10–19. doi: 10.1162/089892903321107783

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746–748.

Mohammed, T., Campbell, R., MacSweeney, M., Milne, E., Hansen, P., & Coleman, M. (2005). Speechreading skill and visual movement sensitivity are related in deaf speechreaders. Perception, 34(2), 205–216.

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. Journal of Neurophysiology, 89(2), 1078–1093. doi: 10.1152/jn.00706.2002

Rosenblum, L. D., Miller, R. M., & Sanchez, K. (2007). Lip-read me now, hear me better later: Cross-modal transfer of talker-familiarity effects. Psychological Science, 18(5), 392–396.

Rosenblum, L. D., Smith, N. M., Nichols, S. M., Lee, J., & Hale, S. (2006). Hearing a face: Cross-modal speaker matching using isolated visible speech. Perception & Psychophysics, 68, 84–93.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. [10.1038/385308a0]. Nature, 385(6614), 308–308.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. Nature, 408(6814), 788. doi: 10.1038/35048669

Soto-Faraco, S., Kingstone, A., & Spence, C. (2003). Multisensory contributions to the perception of motion. Neuropsychologia, 41(13), 1847–1862. doi: 10.1016/s0028-3932(03)00185-4

Spence, C., Senkowski, D., & Roder, B. (2009). Crossmodal processing. [Editorial Introductory]. Exerimental Brain Research, 198(2-3), 107–111. doi: 10.1007/s00221-009-1973–4

Stein, B. E., & Meredith, M. A. (1993). The merging of the senses. Cambridge, MA: The MIT Press.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. [10.1038/nrn2331]. Nature Reviews Neuroscience, 9(4), 255–266.

Sumby, W. H., & Pollack, I. (1954). Visual contribution of speech intelligibility in noise. Journal of the Acoustical Society of America, 26, 212–215.

Vasconcelos, N., Pantoja, J., Belchior, H., Caixeta, F. V., Faber, J., Freire, M. A. M., . . . Ribeiro, S. (2011). Cross-modal responses in the primary visual cortex encode complex objects and correlate with tactile discrimination. Proceedings of the National Academy of Sciences, 108 (37), 15408–15413. doi: 10.1073/pnas.1102780108

Vroomen, J., & De Gelder, B. (2004). Perceptual effects of cross-modal stimulation: Ventriloquism and the freezing phenomenon. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), Handbook of multisensory processes. Cambridge, MA: MIT Press.

18
Eyewitness Testimony and Memory Biases

Cara Laney & Elizabeth F. Loftus

Eyewitnesses can provide very compelling legal testimony, but rather than recording experiences flawlessly, their memories are susceptible to a variety of errors and biases. They (like the rest of us) can make errors in remembering specific details and can even remember whole events that did not actually happen. In this module, we discuss several of the common types of errors, and what they can tell us about human memory and its interactions with the legal system.

Learning Objectives

- Describe the kinds of mistakes that eyewitnesses commonly make and some of the ways that this can impede justice.
- Explain some of the errors that are common in human memory.
- Describe some of the important research that has demonstrated human memory errors and their consequences.

What Is Eyewitness Testimony?

Eyewitness testimony is what happens when a person witnesses a crime (or accident, or other legally important event) and later gets up on the stand and recalls for the court all the details of the witnessed event. It involves a more complicated process than might initially be presumed. It includes what happens during the actual crime to facilitate or hamper witnessing, as well as everything that happens from the time the event is over to the later courtroom appearance. The eyewitness may be interviewed by the police and numerous lawyers, describe the perpetrator to several different people, and make an identification of the perpetrator, among other things.

What can happen to our memory from the time we witness an event to the retelling of that event later? What can influence how we remember, or misremember, highly significant events like a crime or accident? [Image: Robert Couse-Baker, https://goo.gl/OiPUmz, CC BY 2.0, https://goo.gl/BRvSA7]

Why Is Eyewitness Testimony an Important Area of Psychological Research?

When an eyewitness stands up in front of the court and describes what happened from her own perspective, this testimony can be extremely compelling—it is hard for those hearing this testimony to take it "with a grain of salt," or otherwise adjust its power. But to what extent is this necessary?

There is now a wealth of evidence, from research conducted over several decades, suggesting that eyewitness testimony is probably the most persuasive form of evidence presented in court, but in many cases, its accuracy is dubious. There is also evidence that mistaken eyewitness evidence can lead to wrongful conviction—sending people to prison for years or decades, even to death row, for crimes they did not commit. Faulty eyewitness testimony has been implicated in at least 75% of DNA exoneration cases—more than any other cause (Garrett, 2011). In a particularly famous case, a man named Ronald Cotton was identified by a rape victim, Jennifer Thompson, as her rapist, and was found guilty and sentenced to life in prison. After more than 10 years, he was exonerated (and the real rapist identified) based on DNA evidence. For details on this case and other (relatively) lucky individuals whose false convictions were subsequently overturned with DNA evidence, see the Innocence Project website (**http://www.innocenceproject.org/**).

There is also hope, though, that many of the errors may be avoidable if proper precautions are taken during the investigative and judicial processes. Psychological science has taught us what some of those precautions might involve, and we discuss some of that science now.

Misinformation

Misinformation can be introduced into the memory of a witness between the time of seeing an event and reporting it later. Something as straightforward as which sort of traffic sign was in place at an intersection can be confused if subjects are exposed to erroneous information after the initial incident.

In an early study of eyewitness memory, undergraduate subjects first watched a slideshow depicting a small red car driving and then hitting a pedestrian (Loftus, Miller, & Burns, 1978). Some subjects were then asked leading questions about what had happened in the slides. For example, subjects were asked, “How fast was the car traveling when it passed the yield sign?” But this question was actually designed to be misleading, because the original slide included a stop sign rather than a yield sign.

Later, subjects were shown pairs of slides. One of the pair was the original slide containing the stop sign; the other was a replacement slide containing a yield sign. Subjects were asked which of the pair they had previously seen. Subjects who had been asked about the yield sign were likely to pick the slide showing the yield sign, even though they had originally seen the slide with the stop sign. In other words, the misinformation in the leading question led to inaccurate memory.

This phenomenon is called the misinformation effect, because the misinformation that subjects were exposed to after the event (here in the form of a misleading question) apparently contaminates subjects’ memories of what they witnessed. Hundreds of subsequent studies have demonstrated that memory can be contaminated by erroneous information that people are exposed to after they witness an event (see Frenda, Nichols, & Loftus, 2011; Loftus, 2005). The misinformation in these studies has led people to incorrectly remember everything from small but crucial details of a perpetrator's appearance to objects as large as a barn that wasn't there at all.

These studies have demonstrated that young adults (the typical research subjects in psychology) are often susceptible to misinformation, but that children and older adults can be even more susceptible (Bartlett & Memon, 2007; Ceci & Bruck, 1995). In addition, misinformation effects can occur easily, and without any intention to deceive (Allan & Gabbert, 2008). Even slight differences in the wording of a question can lead to misinformation effects. Subjects in one study were more likely to say yes when asked "Did you see the broken headlight?" than when asked "Did you see a broken headlight?" (Loftus, 1975).

Other studies have shown that misinformation can corrupt memory even more easily when it is encountered in social situations (Gabbert, Memon, Allan, & Wright, 2004). This is a problem particularly in cases where more than one person witnesses a crime. In these cases, witnesses tend to talk to one another in the immediate aftermath of the crime, including as they wait for police to arrive. But because different witnesses are different people with different perspectives, they are likely to see or notice different things, and thus remember different things, even when they witness the same event. So when they communicate about the crime later, they not only reinforce common memories for the event, they also contaminate each other's memories for the event (Gabbert, Memon, & Allan, 2003; Paterson & Kemp, 2006; Takarangi, Parker, & Garry, 2006).

The misinformation effect has been modeled in the laboratory. Researchers had subjects watch a video in pairs. Both subjects sat in front of the same screen, but because they wore differently polarized glasses, they saw two different versions of a video, projected onto a screen. So, although they were both watching the same screen, and believed (quite reasonably) that they were watching the same video, they were actually watching two different versions of the video (Garry, French, Kinzett, & Mori, 2008).

In the video, Eric the electrician is seen wandering through an unoccupied house and helping himself to the contents thereof. A total of eight details were different between the two videos. After watching the videos, the "co-witnesses" worked together on 12 memory test questions. Four of these questions dealt with details that were different in the two versions of the video, so subjects had the chance to influence one another. Then subjects worked individually on 20 additional memory test questions. Eight of these were for details that were different in the two videos. Subjects' accuracy was highly dependent on whether they had discussed the details previously. Their accuracy for items they had *not* previously discussed with their co-witness was 79%. But for items that they *had* discussed, their accuracy dropped markedly, to 34%. That is, subjects allowed their co-witnesses to corrupt their memories for what they had seen.

Identifying Perpetrators

In addition to correctly remembering many details of the crimes they witness, eyewitnesses often need to remember the faces and other identifying features of the perpetrators of those crimes. Eyewitnesses are often asked to describe that perpetrator to law enforcement and later to make identifications from books of mug shots or lineups. Here, too, there is a substantial body of research demonstrating that eyewitnesses can make serious, but often understandable and even predictable, errors (Caputo & Dunning, 2007; Cutler & Penrod, 1995).

In most jurisdictions in the United States, lineups are typically conducted with pictures, called **photo spreads**, rather than with actual people standing behind one-way glass (Wells, Memon, & Penrod, 2006). The eyewitness is given a set of small pictures of perhaps six or eight individuals who are dressed similarly and photographed in similar circumstances. One of these individuals is the police suspect, and the remainder are "**foils**" or "fillers" (people known to be innocent of the particular crime under investigation). If the eyewitness identifies the suspect, then the investigation of that suspect is likely to progress. If a witness identifies a foil or no one, then the police may choose to move their investigation in another direction.

Mistakes in identifying perpetrators can be influenced by a number of factors including poor viewing conditions, too little time to view the perpetrator, or too much delay from time of witnessing to identification.

This process is modeled in laboratory studies of eyewitness identifications. In these studies, research subjects witness a mock crime (often as a short video) and then are asked to make an identification from a photo or a live lineup. Sometimes the lineups are target present, meaning that the perpetrator from the mock crime is actually in the lineup, and sometimes they are target absent, meaning that the lineup is made up entirely of foils. The subjects, or **mock witnesses**, are given some instructions and asked to pick the perpetrator out of the lineup. The particular details of the witnessing experience, the instructions, and the lineup members can all influence the extent to which the mock witness is likely to pick the perpetrator out of the lineup, or indeed to make any selection at all. Mock witnesses (and indeed real witnesses) can make errors in two different ways. They can fail to pick the perpetrator out of a target present lineup (by picking a foil or by neglecting to make a selection), or they can pick a foil in a target absent lineup (wherein the only correct choice is to not make a selection).

Some factors have been shown to make eyewitness identification errors particularly likely. These include poor vision or viewing conditions during the crime, particularly stressful witnessing experiences, too little time to view the perpetrator or perpetrators, too much delay between witnessing and identifying, and being asked to identify a perpetrator from a race other than one's own (Bornstein, Deffenbacher, Penrod, & McGorty, 2012; Brigham, Bennett, Meissner, & Mitchell, 2007; Burton, Wilson, Cowan, & Bruce, 1999; Deffenbacher, Bornstein, Penrod, & McGorty, 2004).

It is hard for the legal system to do much about most of these problems. But there are some things that the justice system can do to help lineup identifications "go right." For example, investigators can put together high-quality, fair lineups. A fair lineup is one in which the suspect and each of the foils is equally likely to be chosen by someone who has read an eyewitness description of the perpetrator but who did not actually witness the crime (Brigham, Ready, & Spier, 1990). This means that no one in the lineup should "stick out," and that everyone should match the description given by the eyewitness. Other important recommendations that have come out of this research include better ways to conduct lineups, "double blind" lineups, unbiased instructions for witnesses, and conducting lineups in a sequential fashion (see Technical Working Group for Eyewitness Evidence, 1999; Wells et al., 1998; Wells & Olson, 2003).

Kinds of Memory Biases

Memory is also susceptible to a wide variety of other biases and errors. People can forget events that happened to them and people they once knew. They can mix up details across time and place. They can even remember whole complex events that never happened at all. Importantly, these errors, once made, can be very hard to unmake. A memory is no less "memorable" just because it is wrong.

Some small memory errors are commonplace, and you have no doubt experienced many of them. You set down your keys without paying attention, and then cannot find them later when you go to look for them. You try to come up with a person's name but cannot find it, even though you have the sense that it is right at the tip of your tongue (psychologists actually call this the tip-of-the-tongue effect, or TOT) (Brown, 1991).

For most of our experiences schematas are a benefit and help with information overload. However, they may make it difficult or impossible to recall certain details of a situation later. Do you recall the library as it actually was or the library as approximated by your library schemata? [Dan Kleinman, https://goo.gl/07xyDD, CC BY 2.0, https://goo.gl/BRvSA7]

Other sorts of memory biases are more complicated and longer lasting. For example, it turns out that our expectations and beliefs about how the world works can have huge influences on our memories. Because many aspects of our everyday lives are full of redundancies, our memory systems take advantage of the recurring patterns by forming and using **schemata**, or memory templates (Alba & Hasher, 1983; Brewer & Treyens, 1981). Thus, we know to expect that a library will have shelves and tables and librarians, and so we don't have to spend energy noticing these at the time. The result of this lack of attention, however, is that one is likely to remember schema-consistent information (such as tables), and to remember them in a rather generic way, whether or not they were actually present.

False Memory

Some memory errors are so "large" that they almost belong in a class of their own: **false memories**. Back in the early 1990s a pattern emerged whereby people would go into therapy for depression and other everyday problems, but over the course of the therapy develop memories for violent and horrible victimhood (Loftus & Ketcham, 1994). These patients' therapists claimed that the patients were recovering genuine memories of real childhood abuse, buried deep in their minds for years or even decades. But some experimental psychologists believed that the memories were instead likely to be false—created in therapy. These researchers then set out to see whether it would indeed be possible for wholly false memories to be created by procedures similar to those used in these patients' therapy.

In early false memory studies, undergraduate subjects' family members were recruited to provide events from the students' lives. The student subjects were told that the researchers had talked to their family members and learned about four different events from their childhoods. The researchers asked if the now undergraduate students remembered each of these four events—introduced via short hints. The subjects were asked to write about each of the four events in a booklet and then were interviewed two separate times. The trick was that one of the events came from the researchers rather than the family (and the family had actually assured the researchers that this event had *not* happened to the subject). In the first such study, this researcher-introduced event was a story about being lost in a shopping mall and rescued by an older adult. In this study, after just being asked whether they remembered these events occurring on three separate occasions, a quarter of subjects came to believe that they had indeed been lost in the mall (Loftus & Pickrell, 1995). In subsequent studies, similar procedures were used to get subjects to believe that they nearly drowned and had been rescued by a lifeguard, or that they had spilled punch on the bride's parents at a family wedding, or that they had been attacked by a vicious animal as a child, among other events (Heaps & Nash, 1999; Hyman, Husband, & Billings, 1995; Porter, Yuille, & Lehman, 1999).

More recent false memory studies have used a variety of different manipulations to produce false memories in substantial minorities and even occasional majorities of manipulated subjects (Braun, Ellis, & Loftus, 2002; Lindsay, Hagen, Read, Wade, & Garry, 2004; Mazzoni, Loftus, Seitz, & Lynn, 1999; Seamon, Philbin, & Harrison, 2006; Wade, Garry, Read, & Lindsay, 2002). For example, one group of researchers used a mock-advertising study, wherein subjects were asked to review (fake) advertisements for Disney vacations, to convince subjects that they had once met the character Bugs Bunny at Disneyland—an impossible false memory because Bugs is a Warner Brothers character (Braun et al., 2002). Another group of researchers photoshopped childhood photographs of their subjects into a hot air balloon picture and then asked the subjects to try to remember and describe their hot air balloon experience (Wade et al., 2002). Other researchers gave subjects unmanipulated class photographs from their childhoods along with a fake story about a class prank, and thus enhanced the likelihood that subjects would falsely remember the prank (Lindsay et al., 2004).

Using a false feedback manipulation, we have been able to persuade subjects to falsely remember having a variety of childhood experiences. In these studies, subjects are told (falsely) that a powerful computer system has analyzed questionnaires that they completed previously and has concluded that they had a particular experience years earlier. Subjects apparently believe what the computer says about them and adjust their memories to match this new information. A variety of different false memories have been implanted in this way. In some studies, subjects are told they once got sick on a particular food (Bernstein, Laney, Morris, & Loftus, 2005). These memories can then spill out into other aspects of subjects' lives, such that they often become less interested in eating that food in the future (Bernstein & Loftus, 2009b). Other false memories implanted with this methodology include having an unpleasant experience with the character Pluto at Disneyland and witnessing physical violence between one's parents (Berkowitz, Laney, Morris, Garry, & Loftus, 2008; Laney & Loftus, 2008).

Importantly, once these false memories are implanted—whether through complex methods or simple ones—it is extremely difficult to tell them apart from true memories (Bernstein & Loftus, 2009a; Laney & Loftus, 2008).

Conclusion

To conclude, eyewitness testimony is very powerful and convincing to jurors, even though it is not particularly reliable. Identification errors occur, and these errors can lead to people being falsely accused and even convicted. Likewise, eyewitness memory can be corrupted by leading questions, misinterpretations of events, conversations with co-witnesses, and their own expectations for what should have happened. People can even come to remember whole events that never occurred.

The problems with memory in the legal system are real. But what can we do to start to fix them? A number of specific recommendations have already been made, and many of these are in the process of being implemented (e.g., Steblay & Loftus, 2012; Technical Working Group for Eyewitness Evidence, 1999; Wells et al., 1998). Some of these recommendations are aimed at specific legal procedures, including when and how witnesses should be interviewed, and how lineups should be constructed and conducted. Other recommendations call for appropriate education (often in the form of expert witness testimony) to be provided to jury members and others tasked with assessing eyewitness memory. Eyewitness testimony can be of great value to the legal system, but decades of research now argues that this testimony is often given far more weight than its accuracy justifies.

Outside Resources

Video 1: Eureka Foong's - The Misinformation Effect. This is a student-made video illustrating this phenomenon of altered memory. It was one of the winning entries in the 2014 Noba Student Video Award.
https://www.youtube.com/watch?v=iMPIWkFtd88

Video 2: Ang Rui Xia & Ong Jun Hao's - The Misinformation Effect. Another student-made video exploring the misinformation effect. Also an award winner from 2014.
https://www.youtube.com/watch?v=gsn9iKmOJLQ

Discussion Questions

1. Imagine that you are a juror in a murder case where an eyewitness testifies. In what ways might your knowledge of memory errors affect your use of this testimony?
2. How true to life do you think television shows such as CSI or Law & Order are in their portrayals of eyewitnesses?
3. Many jurisdictions in the United States use "show-ups," where an eyewitness is brought to a suspect (who may be standing on the street or in handcuffs in the back of a police car) and asked, "Is this the perpetrator?" Is this a good or bad idea, from a psychological perspective? Why?

Vocabulary

False memories
Memory for an event that never actually occurred, implanted by experimental manipulation or other means.

Foils
Any member of a lineup (whether live or photograph) other than the suspect.

Misinformation effect
A memory error caused by exposure to incorrect information between the original event (e.g., a crime) and later memory test (e.g., an interview, lineup, or day in court).

Mock witnesses
A research subject who plays the part of a witness in a study.

Photo spreads
A selection of normally small photographs of faces given to a witness for the purpose of identifying a perpetrator.

Schema (plural: schemata)
A memory template, created through repeated exposure to a particular class of objects or events.

References

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*, 203–231.

Allan, K., & Gabbert, F. (2008). I still think it was a banana: Memorable ‘lies’ and forgettable ‘truths’. *Acta Psychologica, 127*(2), 299-308.

Bartlett, J. C., & Memon, A. (2007). Eyewitness memory in young and older adults In: M.P.Toglia, J.D. Read, D.F. Ross & R.C.A Lindsay (Eds.), *The handbook of eyewitness psychology: Volume II: Memory for people* (pp. 309-338). Mahwah, NJ: Lawrence Erlbaum.

Berkowitz, S. R., Laney, C., Morris, E. K., Garry, M., & Loftus, E. F. (2008). Pluto behaving badly: False beliefs and their consequences. *American Journal of Psychology, 121*, 643–660

Bernstein, D. M., & Loftus, E. F. (2009b). The consequences of false memories for food preferences and choices. *Perspectives on Psychological Science, 4*, 135–139.

Bernstein, D. M., & Loftus, E. F., (2009a). How to tell if a particular memory is true or false. *Perspectives on Psychological Science, 4*, 370–374.

Bernstein, D. M., Laney, C., Morris, E. K., & Loftus, E. F. (2005). False memories about food can lead to food avoidance. *Social Cognition, 23*, 11–34.

Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, E. K. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: A meta-analysis of two variables associated with initial memory strength. *Psychology, Crime, & Law, 18*, 473–490.

Braun, K. A., Ellis, R., & Loftus, E. F. (2002). Make my memory: How advertising can change our memories of the past. *Psychology and Marketing, 19*, 1–23.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology, 13*, 207–230.

Brigham, J. C., Bennett, L. B., Meissner, C. A., & Mitchell, T. L. (2007). The influence of race on eyewitness memory. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology, Vol. 2: Memory for people* (pp. 257–281). Mahwah, NJ: Lawrence Erlbaum.

Brown, A. S. (1991). A review of tip of the tongue experience. *Psychological Bulletin, 109*, 79–91.

Burton, A. M., Wilson, S., Cowan, M., & Bruce, V. (1999). Face recognition in poor-quality video: Evidence from security surveillance. *Psychological Science, 10*(3), 243-248.

Caputo, D. D., & Dunning, D. (2007). Distinguishing accurate identifications from erroneous ones: Post-dictive indicators of eyewitness accuracy. In R.C.L. Lindsay, D.F. Ross, J.D. Read, & M.P. Toglia (Eds.), *Handbook of Eyewitness Psychology, Vol 2: Memory for people* (pp.427-449). Mahwah, NJ: Lawrence Erlbaum.

Ceci, S. J., & Bruck, M. (1995). *Jeopardy in the courtroom: A scientific analysis of children\'s testimony*. American Psychological Association.

Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: The eyewitness, psychology and the law.* Cambridge, UK: Cambridge University Press.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687–706.

Frenda, S. J., Nichols, R. M., & Loftus, E. F. (2011). Current issues and advances in misinformation research. *Current Directions in Psychological Science, 20*(1), 20-23.

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology, 9*(2), 215-227.

Garrett, B. L. (2011). *Convicting the innocent*. Cambridge, MA: Harvard University Press.

Garry, M., French, L., Kinzett, T., & Mori, K. (2008). Eyewitness memory following discussion: Using the MORI technique with a Western sample. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 22*(4), 431-439.

Heaps, C., & Nash, M. (1999). Individual differences in imagination inflation. *Psychonomic Bulletin and Review, 6*, 313–138.

Hyman, I. E., Jr., Husband, T. H., & Billings, F. J. (1995). False memories of childhood experiences. *Applied Cognitive Psychology, 9*, 181–197.

Laney, C., & Loftus, E. F. (2008). Emotional content of true and false memories. *Memory, 16*, 500–516.

Lindsay, D. S., Hagen, L., Read, J. D., Wade, K. A., & Garry, M. (2004). True photographs and false memories. *Psychological Science, 15*, 149–154.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361-366.

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7*(4), 560-572.

Loftus, E. F., & Pickrell, J. E. (1995). The formation of false memories. *Psychiatric Annals, 25*, 720–725.

Loftus, E. F., Ketcham, K. (1994). *The myth of repressed memory*. New York, NY: St. Martin’s Press.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19.

Mazzoni, G. A. L., Loftus, E. F., Seitz, A., & Lynn, S.J. (1999). Changing beliefs and memories through dream interpretation. *Applied Cognitive Psychology, 13*, 125–144.

Paterson, H. M., & Kemp, R. I. (2006). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law, 12*(2), 181-191.

Porter, S., Yuille, J. C., & Lehman, D. R. (1999). The nature of real, implanted, and fabricated memories for emotional childhood events: Implications for the recovered memory debate. *Law and Human Behavior, 23*, 517–537.

Seamon, J. G., Philbin, M. M., & Harrison, L. G. (2006). Do you remember proposing marriage to the Pepsi machine? False recollections from a campus walk. *Psychonomic Bulletin & Review, 13*, 752–7596.

Steblay, N. M., & Loftus, E. F. (2012). Eyewitness memory and the legal system. In E. Shafir (Ed.), *The behavioural foundations of public policy* (pp. 145–162). Princeton, NJ: Princeton University Press.

Takarangi, M. K., Parker, S., & Garry, M. (2006). Modernising the misinformation effect: The development of a new stimulus set. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 20*(5), 583-590.

Technical Working Group for Eyewitness Evidence. (1999). *Eyewitness evidence: A trainer\\\\\\\ 's manual for law enforcement*. Research Report. Washington, DC: U.S. Department of Justice.

Wade, K. A., Garry, M., Read, J. D., & Lindsay, S. A. (2002). A picture is worth a thousand lies. *Psychonomic Bulletin and Review, 9*, 597–603.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54*, 277–295.

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7*(2), 45-75.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. A. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human Behavior, 22*, 603–647.

Learning and Memory

19

Factors Influencing Learning

Aaron Benjamin

Learning is a complex process that defies easy definition and description. This module reviews some of the philosophical issues involved with defining learning and describes in some detail the characteristics of learners and of encoding activities that seem to affect how well people can acquire new memories, knowledge, or skills. At the end, we consider a few basic principles that guide whether a particular attempt at learning will be successful or not.

Learning Objectives

- Consider what kinds of activities constitute learning.
- Name multiple forms of learning.
- List some individual differences that affect learning.
- Describe the effect of various encoding activities on learning.
- Describe three general principles of learning.

Introduction

What do you do when studying for an exam? Do you read your class notes and textbook (hopefully not for the very first time)? Do you try to find a quiet place without distraction? Do you use flash cards to test your knowledge? The choices you make reveal your theory of learning, but there is no reason for you to limit yourself to your own intuitions. There is a vast and vibrant science of learning, in which researchers from psychology, education, and neuroscience study basic principles of learning and memory.

When you study for a test, you incorporate your past knowledge into learning this new knowledge. That is, depending on your previous experiences, you will "learn" the material in different ways. [Image: UBC Learning Commons, https://goo.gl/eT0jvd, CC BY 2.0, https://goo.gl/BRvSA7]

In fact, learning is a much broader domain than you might think. Consider: Is listening to music a form of learning? More often, it seems listening to music is a way of avoiding learning. But we know that your brain's response to auditory information changes with your experience with that information, a form of learning called auditory **perceptual learning** (Polley, Steinberg, & Merzenich, 2006). Each time we listen to a song, we hear it differently because of our experience. When we exhibit changes in behavior without having intended to learn something, that is called **implicit learning** (Seger, 1994), and when we exhibit changes in our behavior that reveal the influence of past experience even though we are not attempting to use that experience, that is called **implicit memory** (Richardson-Klavehn & Bjork, 1988).

Other well-studied forms of learning include the types of learning that are general across species. We can't ask a slug to learn a poem or a lemur to learn to bat left-handed, but we can assess learning in other ways. For example, we can look for a change in our responses to things when we are repeatedly stimulated. If you live in a house with a grandfather clock, you know that what was once an annoying and intrusive sound is now probably barely audible to you. Similarly, poking an earthworm again and again is likely to lead to a reduction in its retraction from your touch. These phenomena are forms of **nonassociative learning**, in which single repeated exposure leads to a change in behavior (Pinsker, Kupfermann, Castelluci, & Kandel, 1970). When our response lessens with exposure, it is called **habituation**, and when it increases (like it might with a particularly annoying laugh), it is called **sensitization**. Animals can also learn about relationships between things, such as when an alley cat learns that the sound of janitors working in a restaurant precedes the dumping of delicious new garbage (an example of stimulus-stimulus learning called **classical conditioning**), or when a dog learns to roll over to get a treat (a form of stimulus-response learning called **operant conditioning**). These forms of learning will be covered in the module on Conditioning and Learning (http://noba.to/ajxhcqdr).

Here, we'll review some of the conditions that affect learning, with an eye toward the type of explicit learning we do when trying to learn something. Jenkins (1979) classified experiments on learning and memory into four groups of factors (renamed here): learners, encoding activities, materials, and retrieval. In this module, we'll focus on the first two categories; the module on Memory (http://noba.to/bdc4uger) will consider other factors more generally.

Learners

People bring numerous individual differences with them into memory experiments, and many of these variables affect learning. In the classroom, motivation matters (Pintrich, 2003), though experimental attempts to induce motivation with money yield only modest benefits (Heyer & O'Kelly, 1949). Learners are, however, quite able to allocate more effort to learning prioritized over unimportant materials (Castel, Benjamin, Craik, & Watkins, 2002).

In addition, the organization and planning skills that a learner exhibits matter a lot (Garavalia & Gredler, 2002), suggesting that the efficiency with which one organizes self-guided learning is an important component of learning. We will return to this topic soon.

One well-studied and important variable is working memory capacity. Working memory describes the form of memory we use to hold onto information temporarily. Working memory is used, for example, to keep track of where we are in the course of a complicated math problem, and what the relevant outcomes of prior steps in that problem are. Higher scores on working memory measures are predictive of better reasoning skills (Kyllonen & Christal, 1990), reading comprehension (Daneman & Carpenter, 1980), and even better control of attention (Kane, Conway, Hambrick, & Engle, 2008).

Research attests that we can hold between 5 and 9 individual pieces of information in our working memory at once. This is partly why in the 1950s Bell Labs developed a 7-digit phone number system. [Image: Diamondmagna, https://goo.gl/xeUxfw, CC BY-SA 3.0, https://goo.gl/eLCn2O]

Anxiety also affects the quality of learning. For example, people with math anxiety have a smaller capacity for remembering math-related information in working memory, such as the results of carrying a digit in arithmetic (Ashcraft & Kirk, 2001). Having students write about their specific anxiety seems to reduce the worry associated with tests and increases performance on math tests (Ramirez & Beilock, 2011).

One good place to end this discussion is to consider the role of expertise. Though there probably is a finite capacity on our ability to store information (Landauer, 1986), in practice, this concept is misleading. In fact, because the usual bottleneck to remembering something is our ability to *access* information, not our space to *store* it, having more knowledge or expertise actually enhances our ability to learn new information. A classic example can be seen in comparing a chess master with a chess novice on their ability to learn and remember the positions of pieces on a chessboard (Chase & Simon, 1973). In that experiment, the master remembered the location of many more pieces than the novice, even after only a very short glance. Maybe chess masters are just smarter than the average chess beginner, and have better memory? No: The advantage the expert exhibited only was apparent when the pieces were arranged in a plausible format for an ongoing chess game; when the pieces were placed randomly, both groups did equivalently poorly. Expertise allowed the master to **chunk** (Simon, 1974) multiple pieces into a smaller number of pieces of information—but only when that information was structured in such a way so as to allow the application of that expertise.

Encoding Activities

What we do when we're learning is very important. We've all had the experience of reading something and suddenly coming to the realization that we don't remember a single thing, even the sentence that we just read. *How* we go about **encoding** information determines a lot about how much we remember.

You might think that the most important thing is to *try* to learn. Interestingly, this is not true, at least not completely. Trying to learn a list of words, as compared to just evaluating each word for its part of speech (i.e., noun, verb, adjective) does help you *recall* the words—that is, it helps you remember and write down more of the words later. But it actually impairs your ability to *recognize* the words—to judge on a later list which words are the ones that you studied (Eagle & Leiter, 1964). So this is a case in which **incidental learning**—that is, learning without the intention to learn—is better than **intentional learning**.

Such examples are not particularly rare and are not limited to recognition. Nairne, Pandeirada, and Thompson (2008) showed, for example, that survival processing—thinking about and rating each word in a list for its relevance in a survival scenario—led to much higher recall than intentional learning (and also higher, in fact, than other encoding activities that are also known to lead to high levels of recall). Clearly, merely intending to learn something is not enough. *How* a learner actively processes the material plays a large role; for example, reading words and evaluating their meaning leads to better learning than reading them and evaluating the way that the words look or sound (Craik & Lockhart, 1972). These results suggest that individual differences in motivation will not have a large effect on learning unless learners also have accurate ideas about how to effectively learn material when they care to do so.

Motivation to learn doesn't make much of a difference unless learners use effective strategies for encoding the information they want to retain. Although they're not flashy, methods like spaced practice, interleaving, and frequent testing are among the most effective ways to apply your efforts. [Image: Cali4beach, https://goo.gl/twjIVg, CC BY 2.0, https://goo.gl/BRvSA7]

So, do learners know how to effectively encode material? People allowed to freely allocate their time to study a list of words do remember those words better than a group that doesn't have control over their own study time, though the advantage is relatively small and is limited to the subset of learners who choose to spend more time on the more difficult material (Tullis & Benjamin, 2011). In addition, learners who have an opportunity to review materials that they select for restudy often learn more than another group that is asked to restudy the materials that they *didn't* select for restudy (Kornell & Metcalfe, 2006). However, this advantage also appears to be relatively modest (Kimball, Smith, & Muntean, 2012) and wasn't apparent in a group of older learners (Tullis & Benjamin, 2012). Taken together, all of the evidence seems to support the claim that self-control of learning can be effective, but only when learners have good ideas about what an effective learning strategy is.

One factor that appears to have a big effect and that learners do not always appear to understand is the effect of scheduling repetitions of study. If you are studying for a final exam next week and plan to spend a total of five hours, what is the best way to distribute your study? The evidence is clear that *spacing* one's repetitions apart in time is superior than *massing* them all together (Baddeley & Longman, 1978; Bahrick, Bahrick, Bahrick, & Bahrick, 1993; Melton, 1967). Increasing the spacing between consecutive presentations appears to benefit learning yet further (Landauer & Bjork, 1978).

A similar advantage is evident for the practice of interleaving multiple skills to be learned: For example, baseball batters improved more when they faced a mix of different types of pitches than when they faced the same pitches blocked by type (Hall, Domingues, & Cavazos, 1994).

Students also showed better performance on a test when different types of mathematics problems were interleaved rather than blocked during learning (Taylor & Rohrer, 2010).

One final factor that merits discussion is the role of testing. Educators and students often think about testing as a way of assessing knowledge, and this is indeed an important use of tests. But tests themselves affect memory, because retrieval is one of the most powerful ways of enhancing learning (Roediger & Butler, 2013). Self-testing is an underutilized and potent means of making learning more durable.

General Principles of Learning

We've only begun to scratch the surface here of the many variables that affect the quality and content of learning (Mullin, Herrmann, & Searleman, 1993). But even within this brief examination of the differences between people and the activities they engage in can we see some basic principles of the learning process.

The value of effective metacognition

To be able to guide our own learning effectively, we must be able to evaluate the progress of our learning accurately and choose activities that enhance learning efficiently. It is of little use to study for a long time if a student cannot discern between what material she has or has not mastered, and if additional study activities move her no closer to mastery. **Metacognition** describes the knowledge and skills people have in monitoring and controlling their own learning and memory. We can work to acquire better metacognition by paying attention to our successes and failures in estimating what we do and don't know, and by using testing often to monitor our progress.

Transfer-appropriate processing

Sometimes, it doesn't make sense to talk about whether a particular encoding activity is good or bad for learning. Rather, we can talk about whether that activity is good for learning *as revealed by a particular test*. For example, although reading words for meaning leads to better performance on a test of recall or recognition than paying attention to the pronunciation of the word, it leads to *worse* performance on a test that taps knowledge of that pronunciation, such as whether a previously studied word rhymes with another word (Morris, Bransford, & Franks, 1977). The principle of **transfer-appropriate processing** states that memory is "better" when the test taps the same type of knowledge as the original encoding activity. When thinking about how to learn material, we should always be thinking about the situations in which we are likely to need access to that material. An emergency responder who needs access to learned procedures under conditions of great stress should learn differently from a hobbyist learning to use a new digital camera.

The value of forgetting

In order to not forget things, we employ a variety of tricks (like scribbling a quick note on your hand). However, if we were unable to forget information, it would interfere with learning new or contradictory material. [Image: Andrea Maria Cannata, https://goo.gl/ylTbGG, CC BY-NC 2.0, https://goo.gl/qOP7mj]

Forgetting is sometimes seen as the enemy of learning, but, in fact, forgetting is a highly desirable part of the learning process. The main bottleneck we face in using our knowledge is being able to access it. We have all had the experience of retrieval failure—that is, not being able to remember a piece of information that we know we have, and that we can access easily once the right set of cues is provided. Because access is difficult, it is important to jettison information that is not needed—that is, to forget it. Without forgetting, our minds would become cluttered with out-of-date or irrelevant information. And, just imagine how complicated life would be if we were unable to forget the names of past acquaintances, teachers, or romantic partners.

But the value of forgetting is even greater than that. There is lots of evidence that *some* forgetting is a prerequisite for *more* learning. For example, the previously discussed benefits of distributing practice opportunities may arise in part because of the greater forgetting that takes places between those spaced learning events. It is for this reason that some encoding activities that are difficult and lead to the appearance of slow learning actually lead to superior learning in the long run (Bjork, 2011). When we opt for learning activities that enhance learning quickly, we must be aware that these are not always the same techniques that lead to durable, long-term learning.

Conclusion

To wrap things up, let's think back to the questions we began the module with. What might you now do differently when preparing for an exam? Hopefully, you will think about testing yourself frequently, developing an accurate sense of what you do and do not know, how you are likely to use the knowledge, and using the scheduling of tasks to your advantage. If you are learning a new skill or new material, using the scientific study of learning as a basis for the study and practice decisions you make is a good bet.

Outside Resources

Video: The First 20 hours – How to Learn Anything - Watch a video by Josh Kaufman about how we can get really good at almost anything with 20 hours of efficient practice.
https://www.youtube.com/watch?v=5MgBikgcWnY

Video: The Learning Scientists - Terrific YouTube Channel with videos covering such important topics as interleaving, spaced repetition, and retrieval practice.
https://www.youtube.com/channel/UCjbAmxL6GZXiaoXuNE7cIYg

Video: What we learn before we're born - In this video, science writer Annie Murphy Paul answers the question "When does learning begin?" She covers through new research that shows how much we learn in the womb — from the lilt of our native language to our soon-to-be-favorite foods.
https://www.ted.com/talks/annie_murphy_paul_what_we_learn_before_we_re_born

Web: Neuroscience News - This is a science website dedicated to neuroscience research, with this page addressing fascinating new memory research.
http://neurosciencenews.com/neuroscience-terms/memory-research/

Web: The Learning Scientists - A websitecreated by three psychologists who wanted to make scientific research on learning more accessible to students, teachers, and other educators.
http://www.learningscientists.org/

Discussion Questions

1. How would you best design a computer program to help someone learn a new foreign language? Think about some of the principles of learning outlined in this module and how those principles could be instantiated in "rules" in a computer program.
2. Would you rather have a really good memory or really good metacognition? How might you train someone to develop better metacognition if he or she doesn't have a very good memory, and what would be the consequences of that training?
3. In what kinds of situations not discussed here might you find a benefit of forgetting on learning?

Vocabulary

Chunk
The process of grouping information together using our knowledge.

Classical conditioning
Describes stimulus-stimulus associative learning.

Encoding
The pact of putting information into memory.

Habituation
Occurs when the response to a stimulus decreases with exposure.

Implicit learning
Occurs when we acquire information without intent that we cannot easily express.

Implicit memory
A type of long-term memory that does not require conscious thought to encode. It's the type of memory one makes without intent.

Incidental learning
Any type of learning that happens without the intention to learn.

Intentional learning
Any type of learning that happens when motivated by intention.

Metacognition
Describes the knowledge and skills people have in monitoring and controlling their own learning and memory.

Nonassociative learning
Occurs when a single repeated exposure leads to a change in behavior.

Operant conditioning
Describes stimulus-response associative learning.

Perceptual learning

Occurs when aspects of our perception changes as a function of experience.

Sensitization
Occurs when the response to a stimulus increases with exposure

Transfer-appropriate processing
A principle that states that memory performance is superior when a test taps the same cognitive processes as the original encoding activity.

Working memory
The form of memory we use to hold onto information temporarily, usually for the purposes of manipulation.

References

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology*: General, 130, 224–237.

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics*, 21, 627–635.

Bahrick, H. P., Bahrick, L. E., Bahrick, A. S., & Bahrick, P. O. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, 4, 316–321.

Bjork, R. A. (2011). On the symbiosis of learning, remembering, and forgetting. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 1–22). London, UK: Psychology Press.

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition*, 30, 1078–1085.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.

Eagle, M., & Leiter, E. (1964). Recall and recognition in intentional and incidental learning. *Journal of Experimental Psychology*, 68, 58–63.

Garavalia, L. S., & Gredler, M. E. (2002). Prior achievement, aptitude, and use of learning strategies as predictors of college student achievement. *College Student Journal*, 36, 616–626.

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills*, 78, 835–841.

Heyer, A. W., Jr., & O'Kelly, L. I. (1949). Studies in motivation and retention: II. Retention of nonsense syllables learned under different degrees of motivation. *Journal of Psychology: Interdisciplinary and Applied*, 27, 143–152.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing and human memory* (pp. 429–446). Hillsdale, NJ: Erlbaum.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2008). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in Working Memory* (pp. 22–48). New York, NY: Oxford University Press.

Kimball, D. R., Smith, T. A., & Muntean, W. J. (2012). Does delaying judgments of learning really improve the efficacy of study decisions? Not so much. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 923–954.

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32, 609–622.*

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity. *Intelligence*, 14, 389–433.

Landauer, T. K. (1986). How much do people remember? Some estimates of the quantity of learned information in long-term memory. *Cognitive Science, 10*, 477–493.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Melton, A. W. (1967). Repetition and retrieval from memory. *Science*, 158, 532.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 519–533.

Mullin, P. A., Herrmann, D. J., & Searleman, A. (1993). Forgotten variables in memory theory and research. *Memory*, 1, 43–64.

Nairne, J. S., Pandeirada, J. N. S., & Thompson, S. R. (2008). Adaptive memory: the comparative value of survival processing. *Psychological Science*, 19, 176–180.

Pinsker, H., Kupfermann, I., Castelluci, V., & Kandel, E. (1970). Habituation and dishabituation of the gill-withdrawal reflex in Aplysia. *Science*, 167, 1740–1742.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95, 667–686.

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. *The Journal of Neuroscience*, 26, 4970–4982.

Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science*, 331, 211–213.

Richardson-Klavehn, A. & Bjork, R.A. (1988). Measures of memory. *Annual Review of Psychology*, 39, 475–543.

Roediger, H. L., & Butler, A.C. (2013). Retrieval practice (testing) effect. In H. L. Pashler (Ed.), *Encyclopedia of the mind*. Los Angeles, CA: Sage Publishing Co.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, 115, 163–196.

Simon, H. A. (1974). How big is a chunk? *Science*, 4124, 482–488.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology*, 24, 837–848.

Tullis, J. G., & Benjamin, A. S. (2012). Consequences of restudy choices in younger and older learners. *Psychonomic Bulletin & Review*, 19, 743–749.

Tullis, J. G., & Benjamin, A. S. (2011). On the effectiveness of self-paced learning. *Journal of Memory and Language*, 64, 109–118.

20
Memory (Encoding, Storage, Retrieval)

Kathleen B. McDermott & Henry L. Roediger

"Memory" is a single term that reflects a number of different abilities: holding information briefly while working with it (working memory), remembering episodes of one's life (episodic memory), and our general knowledge of facts of the world (semantic memory), among other types. Remembering episodes involves three processes: encoding information (learning it, by perceiving it and relating it to past knowledge), storing it (maintaining it over time), and then retrieving it (accessing the information when needed). Failures can occur at any stage, leading to forgetting or to having false memories. The key to improving one's memory is to improve processes of encoding and to use techniques that guarantee effective retrieval. Good encoding techniques include relating new information to what one already knows, forming mental images, and creating associations among information that needs to be remembered. The key to good retrieval is developing effective cues that will lead the rememberer back to the encoded information. Classic mnemonic systems, known since the time of the ancient Greeks and still used by some today, can greatly improve one's memory abilities.

Learning Objectives

- Define and note differences between the following forms of memory: working memory, episodic memory, semantic memory, collective memory.
- Describe the three stages in the process of learning and remembering.
- Describe strategies that can be used to enhance the original learning or encoding of information.
- Describe strategies that can improve the process of retrieval.
- Describe why the classic mnemonic device, the method of loci, works so well.

Introduction

In 2013, Simon Reinhard sat in front of 60 people in a room at Washington University, where he memorized an increasingly long series of digits. On the first round, a computer generated 10 random digits—6 1 9 4 8 5 6 3 7 1—on a screen for 10 seconds. After the series disappeared, Simon typed them into his computer. His recollection was perfect. In the next phase, 20 digits appeared on the screen for 20 seconds. Again, Simon got them all correct. No one in the audience (mostly professors, graduate students, and undergraduate students) could recall the 20 digits perfectly. Then came 30 digits, studied for 30 seconds; once again, Simon didn't misplace even a single digit. For a final trial, 50 digits appeared on the screen for 50 seconds, and again, Simon got them all right. In fact, Simon would have been happy to keep going. His record in this task—called "forward digit span"—is 240 digits!

In some ways memory is like file drawers where you store mental information. Memory is also a series of processes: how does that information get filed to begin with and how does it get retrieved when needed? [Image: M Cruz, https://goo.gl/DhOMgp, CC BY-SA 4.0, https://goo.gl/SWjq94]

When most of us witness a performance like that of Simon Reinhard, we think one of two things: First, maybe he's cheating somehow. (No, he is not.) Second, Simon must have abilities more advanced than the rest of humankind. After all, psychologists established many years ago that the normal memory span for adults is about 7 digits, with some of us able to recall a few more and others a few less (Miller, 1956). That is why the first phone numbers were limited to 7 digits—psychologists determined that many errors occurred (costing the phone company money) when the number was increased to even 8 digits. But in normal testing, no one gets 50 digits correct in a row, much less 240. So, does Simon Reinhard simply have a photographic memory? He does not. Instead, Simon has taught himself simple strategies for remembering that have greatly increased his capacity for remembering virtually any type of material—digits, words, faces and names, poetry, historical dates, and so on. Twelve years earlier, before he started training his memory abilities, he had a digit span of 7, just like most of us. Simon has been training his abilities for about 10 years as of this writing, and has risen to be in the top two of “memory athletes.” In 2012, he came in second place in the World Memory Championships (composed of 11 tasks), held in London. He currently ranks second in the world, behind another German competitor, Johannes Mallow. In this module, we reveal what psychologists and others have learned about memory, and we also explain the general principles by which you can improve your own memory for factual material.

Varieties of Memory

To be a good chess player you have to learn to increase working memory so you can plan ahead for several offensive moves while simultaneously anticipating - through use of memory - how the other player could counter each of your planned moves. [Image: karpidis, https://goo.gl/EhzMKM, CC BY-SA 2.0, https://goo.gl/jSSrcO]

For most of us, remembering digits relies on *short-term memory,* or *working memory*—the ability to hold information in our minds for a brief time and work with it (e.g., multiplying 24 x 17 without using paper would rely on working memory). Another type of memory is **episodic memory**—the ability to remember the episodes of our lives. If you were given the task of recalling everything you did 2 days ago, that would be a test of episodic memory; you would be required to mentally travel through the day in your mind and note the main events. **Semantic memory** is our storehouse of more-or-less permanent knowledge, such as the meanings of words in a language (e.g., the meaning of “parasol”) and the huge collection of facts about the world (e.g., there are 196 countries in the world, and 206 bones in your body). *Collective memory* refers to the kind of memory that people in a group share (whether family, community, schoolmates, or citizens of a state or a country). For example, residents of small towns often strongly identify with those towns, remembering the local customs and historical events in a unique way. That is, the community’s collective memory passes stories and recollections between neighbors and to future generations, forming a memory system unto itself.

Psychologists continue to debate the classification of types of memory, as well as which types rely on others (Tulving, 2007), but for this module we will focus on episodic memory. Episodic memory is usually what people think of when they hear the word “memory.” For example, when people say that an older relative is "losing her memory" due to Alzheimer's disease, the type of memory-loss they are referring to is the inability to recall events, or episodic memory. (Semantic memory is actually preserved in early-stage Alzheimer's disease.) Although remembering specific events that have happened over the course of one's entire life (e.g., your experiences in sixth grade) can be referred to as **autobiographical memory**, we will focus primarily on the episodic memories of more recent events.

Three Stages of the Learning/Memory Process

Psychologists distinguish between three necessary stages in the learning and memory process: **encoding**, **storage**, and **retrieval** (Melton, 1963). Encoding is defined as the initial learning of information; storage refers to maintaining information over time; retrieval is the ability to access information when you need it. If you meet someone for the first time at a party, you need to encode her name (Lyn Goff) while you associate her name with her face. Then you need to maintain the information over time. If you see her a week later, you need to recognize her face and have it serve as a cue to retrieve her name. Any successful act of remembering requires that all three stages be intact. However, two types of errors can also occur. Forgetting is one type: you see the person you met at the party and you cannot recall her name. The other error is misremembering (false recall or false recognition): you see someone who looks like Lyn Goff and call the person by that name (false recognition of the face). Or, you might see the real Lyn Goff, recognize her face, but then call her by the name of another woman you met at the party (misrecall of her name).

Whenever forgetting or misremembering occurs, we can ask, at which stage in the learning/memory process was there a failure?—though it is often difficult to answer this question with precision. One reason for this inaccuracy is that the three stages are not as discrete as our description implies. Rather, all three stages depend on one another. How we encode information determines how it will be stored and what cues will be effective when we try to retrieve it. And too, the act of retrieval itself also changes the way information is subsequently remembered, usually aiding later recall of the retrieved information. The central point for now is that the three stages—encoding, storage, and retrieval—affect one another, and are inextricably bound together.

Encoding

Encoding refers to the initial experience of perceiving and learning information. Psychologists often study recall by having participants study a list of pictures or words. Encoding in these situations is fairly straightforward. However, "real life" encoding is much more challenging.

When you walk across campus, for example, you encounter countless sights and sounds—friends passing by, people playing Frisbee, music in the air. The physical and mental environments are much too rich for you to encode all the happenings around you or the internal thoughts you have in response to them. So, an important first principle of encoding is that it is selective: we attend to some events in our environment and we ignore others. A second point about encoding is that it is prolific; we are always encoding the events of our lives—attending to the world, trying to understand it. Normally this presents no problem, as our days are filled with routine occurrences, so we don't need to pay attention to everything. But if something does happen that seems strange—during your daily walk across campus, you see a giraffe—then we pay close attention and try to understand why we are seeing what we are seeing.

Right after your typical walk across campus (one without the appearance of a giraffe), you would be able to remember the events reasonably well if you were asked. You could say whom you bumped into, what song was playing from a radio, and so on. However, suppose someone asked you to recall the same walk a month later. You wouldn't stand a chance. You would likely be able to recount the basics of a typical walk across campus, but not the precise details of that particular walk. Yet, if you had seen a giraffe during that walk, the event would have been fixed in your mind for a long time, probably for the rest of your life. You would tell your friends about it, and, on later occasions when you saw a giraffe, you might be reminded of the day you saw one on campus. Psychologists have long pinpointed **distinctiveness**—having an event stand out as quite different from a background of similar events—as a key to remembering events (Hunt, 2003).

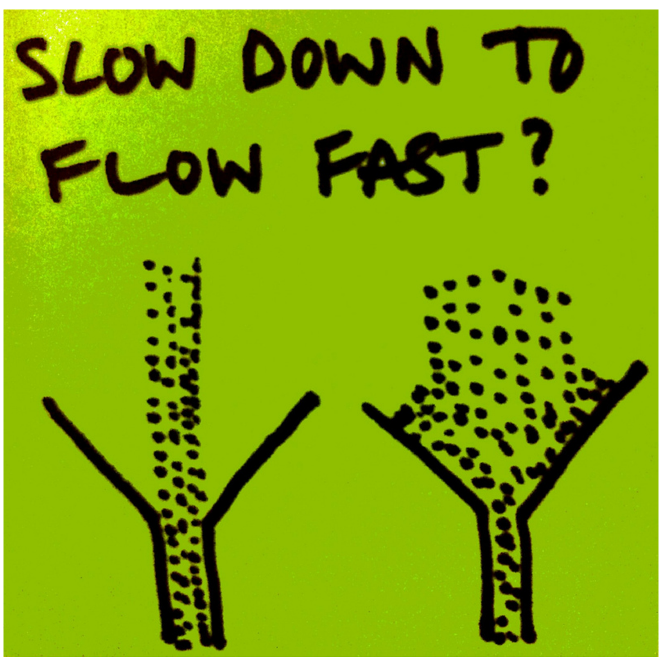
A giraffe in the context of a zoo or its natural habitat may register as nothing more than ordinary, but put it in another setting - in the middle of a campus or a busy city - and its level of distinctiveness increases dramatically. Distinctiveness is a key attribute to remembering events. [Image: Colin J Babb, https://goo.gl/Cci2yl, CC BY-SA 2.0, https://goo.gl/jSSrcO]

In addition, when vivid memories are tinged with strong emotional content, they often seem to leave a permanent mark on us. Public tragedies, such as terrorist attacks, often create vivid memories in those who witnessed them. But even those of us not directly involved in such events may have vivid memories of them, including memories of first hearing about them.

For example, many people are able to recall their exact physical location when they first learned about the assassination or accidental death of a national figure. The term **flashbulb memory** was originally coined by Brown and Kulik (1977) to describe this sort of vivid memory of finding out an important piece of news. The name refers to how some memories seem to be captured in the mind like a flash photograph; because of the distinctiveness and emotionality of the news, they seem to become permanently etched in the mind with exceptional clarity compared to other memories.

Take a moment and think back on your own life. Is there a particular memory that seems sharper than others? A memory where you can recall unusual details, like the colors of mundane things around you, or the exact positions of surrounding objects? Although people have great confidence in flashbulb memories like these, the truth is, our objective accuracy with them is far from perfect (Talarico & Rubin, 2003). That is, even though people may have great confidence in what they recall, their memories are not as accurate (e.g., what the actual colors were; where objects were truly placed) as they tend to imagine. Nonetheless, all other things being equal, distinctive and emotional events are well-remembered.

Details do not leap perfectly from the world into a person's mind. We might say that we went to a party and remember it, but what we remember is (at best) what we encoded. As noted above, the process of encoding is selective, and in complex situations, relatively few of many possible details are noticed and encoded. The process of encoding always involves **recoding**—that is, taking the information from the form it is delivered to us and then converting it in a way that we can make sense of it. For example, you might try to remember the colors of a rainbow by using the acronym ROY G BIV (red, orange, yellow, green, blue, indigo, violet). The process of recoding the colors into a name can help us to remember. However, recoding can also introduce errors—when we accidentally add information during encoding, then remember that *new* material as if it had been part of the actual experience (as discussed below).



Although it requires more effort, using images and associations can improve the process of recoding. [Image: psd, https://goo.gl/9xjcDe, CC BY 2.0, https://goo.gl/9uSnqN]

Psychologists have studied many

recoding strategies that can be used during study to improve retention. First, research advises that, as we study, we should think of the meaning of the events (Craik & Lockhart, 1972), and we should try to relate new events to information we already know. This helps us form associations that we can use to retrieve information later. Second, imagining events also makes them more memorable; creating vivid images out of information (even verbal information) can greatly improve later recall (Bower & Reitman, 1972). Creating imagery is part of the technique Simon Reinhard uses to remember huge numbers of digits, but we can all use images to encode information more effectively. The basic concept behind good encoding strategies is to form distinctive memories (ones that stand out), and to form links or associations among memories to help later retrieval (Hunt & McDaniel, 1993). Using study strategies such as the ones described here is challenging, but the effort is well worth the benefits of enhanced learning and retention.

We emphasized earlier that encoding is selective: people cannot encode all information they are exposed to. However, recoding can add information that was not even seen or heard during the initial encoding phase. Several of the recoding processes, like forming associations between memories, can happen without our awareness. This is one reason people can sometimes remember events that did not actually happen—because during the process of recoding, details got added. One common way of inducing false memories in the laboratory employs a word-list technique (Deese, 1959; Roediger & McDermott, 1995). Participants hear lists of 15 words, like *door, glass, pane, shade, ledge, sill, house, open, curtain, frame, view, breeze, sash, screen,* and *shutter.* Later, participants are given a test in which they are shown a list of words and asked to pick out the ones they'd heard earlier. This second list contains some words from the first list (e.g., *door, pane, frame*) and some words not from the list (e.g., *arm, phone, bottle*). In this example, one of the words on the second list is *window*, which—importantly—does not appear in the first list, but which is related to other words in that list. When subjects were tested with the second list, they were reasonably accurate with the studied words (*door*, etc.), recognizing them 72% of the time. However, when *window* was on the test, they falsely recognized it as having been on the list 84% of the time (Stadler, Roediger, & McDermott, 1999). The same thing happened with many other lists the authors used. This phenomenon is referred to as the DRM (for Deese-Roediger-McDermott) effect. One explanation for such results is that, while students listened to items in the list, the words triggered the students to think about *window*, even though *window* was never presented. In this way, people seem to encode events that are not actually part of their experience.

Because humans are creative, we are always going beyond the information we are given: we automatically make associations and infer from them what is happening. But, as with the word association mix-up above, sometimes we make false memories from our inferences—remembering the inferences themselves as if they were actual experiences. To illustrate this, Brewer (1977) gave people sentences to remember that were designed to elicit *pragmatic inferences*. Inferences, in general, refer to instances when something is not explicitly stated, but we are still able to guess the undisclosed intention. For example, if your friend told you that she didn't want to go out to eat, you may infer that she doesn't have the money to go out, or that she's too tired. With *pragmatic* inferences, there is usually *one* particular inference you're likely to make. Consider the statement Brewer (1977) gave her participants: "The karate champion hit the cinder block." After hearing or seeing this sentence, participants who were given a memory test tended to remember the statement as having been, "The karate champion *broke* the cinder block." This remembered statement is not necessarily a *logical* inference (i. e., it is perfectly reasonable that a karate champion could hit a cinder block without breaking it). Nevertheless, the *pragmatic* conclusion from hearing such a sentence is that the block was likely broken. The participants remembered this inference they made while hearing the sentence in place of the actual words that were in the sentence (see also McDermott & Chan, 2006).

Encoding—the initial registration of information—is essential in the learning and memory process. Unless an event is encoded in some fashion, it will not be successfully remembered later. However, just because an event is encoded (even if it is encoded well), there's no guarantee that it will be remembered later.

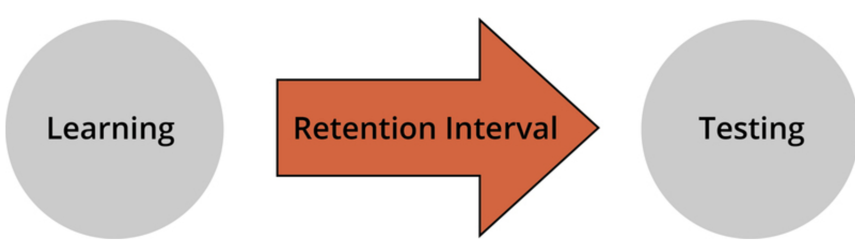
Storage

Every experience we have changes our brains. That may seem like a bold, even strange, claim at first, but it's true. We encode each of our experiences within the structures of the nervous system, making new impressions in the process—and each of those impressions involves changes in the brain. Psychologists (and neurobiologists) say that experiences leave **memory traces**, or **engrams** (the two terms are synonyms). Memories have to be stored somewhere in the brain, so in order to do so, the brain biochemically alters itself and its neural tissue. Just like you might write yourself a note to remind you of something, the brain "writes" a memory trace, changing its own

Memory traces, or engrams, are NOT perfectly preserved recordings of past experiences. The traces are combined with current knowledge to reconstruct what we think happened in the past. [Simon Bierdwald, https://goo.gl/JDhdCE, CC BY-NC-SA 2.0, https://goo.gl/jSSrcO] physical composition to do so. The basic idea is that events (occurrences in our environment) create engrams through a process of **consolidation**: the neural changes that occur after learning to create the memory trace of an experience. Although neurobiologists are concerned with exactly what neural processes change when memories are created, for psychologists, the term *memory trace* simply refers to the physical change in the nervous system (whatever that may be, exactly) that represents our experience.

Although the concept of engram or memory trace is extremely useful, we shouldn't take the term too literally. It is important to understand that memory traces are not perfect little packets of information that lie dormant in the brain, waiting to be called forward to give an accurate report of past experience. Memory traces are not like video or audio recordings, capturing experience with great accuracy; as discussed earlier, we often have errors in our memory, which would not exist if memory traces were perfect packets of information. Thus, it is wrong to think that remembering involves simply "reading out" a faithful record of past experience. Rather, when we remember past events, we reconstruct them with the aid of our memory traces—but also with our current belief of what happened. For example, if you were trying to recall for the police who started a fight at a bar, you may not have a memory trace of who pushed whom first. However, let's say you remember that one of the guys held the door open for you. When thinking back to the start of the fight, this knowledge (of how one guy was friendly to you) may unconsciously influence your memory of what happened in favor of the nice guy. Thus, memory is a construction of what you actually recall and what you believe happened. In a phrase, remembering is reconstructive (we reconstruct our past with the aid of memory traces) not reproductive (a perfect reproduction or recreation of the past).

Psychologists refer to the time between learning and testing as the retention interval. Memories can consolidate during that time, aiding retention. However, experiences can also occur that undermine the memory. For example, think of what you had for lunch yesterday



—a pretty easy task. However, if you had to recall what you had for lunch 17 days ago, you may well fail (assuming you don't eat the same thing every day). The 16 lunches you've had since that one have created **retroactive interference**. Retroactive interference refers to new activities (i.e., the subsequent lunches) during the retention interval (i.e., the time between the lunch 17 days ago and now) that interfere with retrieving the specific, older memory (i.e., the lunch details from 17 days ago). But just as newer things can interfere with remembering older things, so can the opposite happen. *Proactive interference* is when past memories interfere with the encoding of new ones. For example, if you have ever studied a second language, often times the grammar and vocabulary of your native language will pop into your head, impairing your fluency in the foreign language.

Retroactive interference is one of the main causes of forgetting (McGeoch, 1932). In the module *Eyewitness Testimony and Memory Biases* **http://noba.to/uy49tm37** Elizabeth Loftus describes her fascinating work on eyewitness memory, in which she shows how memory for an event can be changed via misinformation supplied during the retention interval. For example, if you witnessed a car crash but subsequently heard people describing it from their own perspective, this new information may interfere with or disrupt your own personal recollection of the crash. In fact, you may even come to remember the event happening exactly as the others described it! This **misinformation effect** in eyewitness memory represents a type of retroactive interference that can occur during the retention interval (see Loftus [2005] for a review). Of course, if correct information is given during the retention interval, the witness's memory will usually be improved.

Although interference may arise between the occurrence of an event and the attempt to recall it, *the effect itself is always expressed when we retrieve memories*, the topic to which we turn next.

Retrieval

Endel Tulving argued that "the key process in memory is retrieval" (1991, p. 91). Why should retrieval be given more prominence than encoding or storage? For one thing, if information were encoded and stored but could not be retrieved, it would be useless. As discussed previously in this module, we encode and store thousands of events—conversations, sights and sounds—every day, creating memory traces. However, we later access only a tiny portion of what we've taken in. Most of our memories will never be used—in the sense of being brought back to mind, consciously. This fact seems so obvious that we rarely reflect on it. All those events that happened to you in the fourth grade that seemed so important then? Now, many years later, you would struggle to remember even a few. You may wonder if the traces of those memories still exist in some latent form. Unfortunately, with currently available methods, it is impossible to know.

Psychologists distinguish information that is available in memory from that which is accessible (Tulving & Pearlstone, 1966). *Available* information is the information that is stored in memory—but precisely how much and what types are stored cannot be known. That is, all we can know is what information we can retrieve—*accessible* information. The assumption is that accessible information represents only a tiny slice of the information available in our brains. Most of us have had the experience of trying to remember some fact or event, giving up, and then—all of a sudden!—it comes to us at a later time, even after we've stopped trying to remember it. Similarly, we all know the experience of failing to recall a fact, but then, if we are given several choices (as in a multiple-choice test), we are easily able to recognize it.

We can't know the entirety of what is in our memory, but only that portion we can actually retrieve. Something that cannot be retrieved now and which is seemingly gone from memory may, with different cues applied, reemerge. [Image: Ores2k, https://goo.gl/1du8Qe, CC BY-NC-SA 2.0, https://goo.gl/jSSrcO]

What factors determine what information can be retrieved from memory? One critical factor is the type of hints, or *cues*, in the environment. You may hear a song on the radio that suddenly evokes memories of an earlier time in your life, even if you were not trying to remember it when the song came on. Nevertheless, the song is closely associated with that time, so it brings the experience to mind.

The general principle that underlies the effectiveness of retrieval cues is the **encoding specificity principle** (Tulving & Thomson, 1973): when people encode information, they do so in specific ways. For example, take the song on the radio: perhaps you heard it while you were at a terrific party, having a great, philosophical conversation with a friend. Thus, the song became part of that whole complex experience. Years later, even though you haven't thought about that party in ages, when you hear the song on the radio, the whole experience rushes back to you. In general, the encoding specificity principle states that, to the extent a retrieval cue (the song) matches or overlaps the memory trace of an experience (the party, the conversation), it will be effective in evoking the memory. A classic experiment on the encoding specificity principle had participants memorize a set of words in a unique setting. Later, the participants were tested on the word sets, either in the same location they learned the words or a different one. As a result of encoding specificity, the students who took the test in the same place they learned the words were actually able to recall more words (Godden & Baddeley, 1975) than the students who took the test in a new setting.

One caution with this principle, though, is that, for the cue to work, it can't match too many other experiences (Nairne, 2002; Watkins, 1975). Consider a lab experiment. Suppose you study 100 items; 99 are words, and one is a picture—of a penguin, item 50 in the list. Afterwards, the cue "recall the picture" would evoke "penguin" perfectly. No one would miss it. However, if the *word* "penguin" were placed in the same spot among the other 99 words, its memorability would be exceptionally worse. This outcome shows the power of distinctiveness that we discussed in the section on encoding: one picture is perfectly recalled from among 99 words because it stands out. Now consider what would happen if the experiment were repeated, but there were 25 pictures distributed within the 100-item list. Although the picture of the penguin would still be there, the probability that the cue "recall the picture" (at item 50) would be useful for the penguin would drop correspondingly. Watkins (1975) referred to this outcome as demonstrating the **cue overload principle**. That is, to be effective, a retrieval cue cannot be overloaded with too many memories. For the cue "recall the picture" to be effective, it should only match one item in the target set (as in the one-picture, 99-word case).

To sum up how memory cues function: for a retrieval cue to be effective, a match must exist between the cue and the desired target memory; furthermore, to produce the best retrieval, the cue-target relationship should be distinctive. Next, we will see how the encoding specificity principle can work in practice.

Psychologists measure memory performance by using production tests (involving recall) or recognition tests (involving the selection of correct from incorrect information, e.g., a multiple-choice test). For example, with our list of 100 words, one group of people might be asked to recall the list in any order (a free recall test), while a different group might be asked to circle the 100 studied words out of a mix with another 100, unstudied words (a recognition test). In this situation, the recognition test would likely produce better performance from participants than the recall test.

We usually think of recognition tests as being quite easy, because the cue for retrieval is a copy of the actual event that was presented for study. After all, what could be a better cue than the exact target (memory) the person is trying to access? In most cases, this line of reasoning is true; nevertheless, recognition tests do not provide perfect indexes of what is stored in memory. That is, you can fail to recognize a target staring you right in the face, yet be able to recall it later with a different set of cues (Watkins & Tulving, 1975). For example, suppose you had the task of recognizing the surnames of famous authors. At first, you might think that being given the actual last name would always be the best cue. However, research has shown this not necessarily to be true (Muter, 1984). When given names such as Tolstoy, Shaw, Shakespeare, and Lee, subjects might well say that Tolstoy and Shakespeare are famous authors, whereas Shaw and Lee are not. But, when given a cued recall test using first names, people often recall items (produce them) that they had failed to recognize before. For example, in this instance, a cue like *George Bernard* _______ often leads to a recall of "Shaw," even though people initially failed to recognize *Shaw* as a famous author's name. Yet, when given the cue "William," people may not come up with Shakespeare, because William is a common name that matches many people (the cue overload principle at work). This strange fact—that recall can sometimes lead to better performance than recognition—can be explained by the encoding specificity principle. As a cue, *George Bernard* _________ matches the way the famous writer is stored in memory better than does his surname, Shaw, does (even though it is the target). Further, the match is quite distinctive with *George Bernard* ___________, but the cue *William* _________________ is much more overloaded (Prince William, William Yeats, William Faulkner, will.i.am).

The phenomenon we have been describing is called the *recognition failure of recallable words*, which highlights the point that a cue will be most effective depending on how the information has been encoded (Tulving & Thomson, 1973). The point is, the cues that work best to evoke retrieval are those that recreate the event or name to be remembered, whereas sometimes even the target itself, such as *Shaw* in the above example, is not the best cue. Which cue will be most effective depends on how the information has been encoded.

Whenever we think about our past, we engage in the act of retrieval. We usually think that retrieval is an objective act because we tend to imagine that retrieving a memory is like pulling a book from a shelf, and after we are done with it, we return the book to the shelf just as it was. However, research shows this assumption to be false; far from being a static repository of data, the memory is constantly changing. In fact, every time we retrieve a memory, it is altered. For example, the act of retrieval itself (of a fact, concept, or event) makes the retrieved memory much more likely to be retrieved again, a phenomenon called the *testing effect* or the *retrieval practice effect* (Pyc & Rawson, 2009; Roediger & Karpicke, 2006). However, retrieving some information can actually cause us to forget other information related to it, a phenomenon called *retrieval-induced forgetting* (Anderson, Bjork, & Bjork, 1994). Thus the act of retrieval can be a double-edged sword—strengthening the memory just retrieved (usually by a large amount) but harming related information (though this effect is often relatively small).

As discussed earlier, retrieval of distant memories is reconstructive. We weave the concrete bits and pieces of events in with assumptions and preferences to form a coherent story (Bartlett, 1932). For example, if during your 10th birthday, your dog got to your cake before you did, you would likely tell that story for years afterward. Say, then, in later years you misremember where the dog actually found the cake, but repeat that error over and over during subsequent retellings of the story. Over time, that inaccuracy would become a basic fact of the event in your mind. Just as retrieval practice (repetition) enhances accurate memories, so will it strengthen errors or false memories (McDermott, 2006). Sometimes memories can even be manufactured just from hearing a vivid story. Consider the following episode, recounted by Jean Piaget, the famous developmental psychologist, from his childhood:

> One of my first memories would date, if it were true, from my second year. I can still see, most clearly, the following scene, in which I believed until I was about 15. I was sitting in my pram . . . when a man tried to kidnap me. I was held in by the strap fastened round me while my nurse bravely tried to stand between me and the thief. She received various scratches, and I can still vaguely see those on her face. . . . When I was about 15, my parents received a letter from my former nurse saying that she had been converted to the Salvation Army. She wanted to confess her past faults, and in particular to return the watch she had been given as a reward on this occasion. She had made up the whole story, faking the scratches. I therefore must have heard, as a child, this story, which my parents believed, and projected it into the past in the form of a visual memory. . . . Many real memories are doubtless of the same order. (Norman & Schacter, 1997, pp. 187–188)

Piaget's vivid account represents a case of a pure reconstructive memory. He heard the tale told repeatedly, and doubtless told it (and thought about it) himself. The repeated telling cemented the events as though they had really happened, just as we are all open to the possibility of having "many real memories ... of the same order." The fact that one can remember precise details (the location, the scratches) does not necessarily indicate that the memory is true, a point that has been confirmed in laboratory studies, too (e.g., Norman & Schacter, 1997).

Putting It All Together: Improving Your Memory

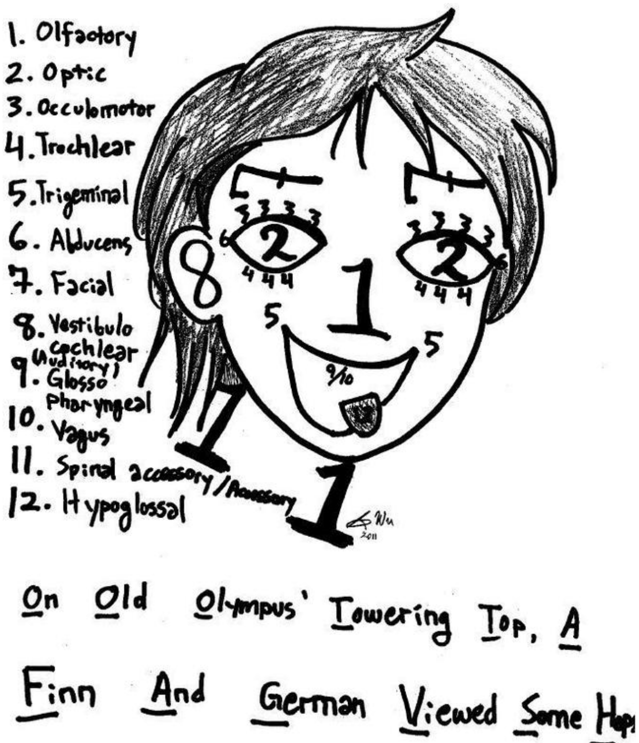
A central theme of this module has been the importance of the encoding and retrieval processes, and their interaction. To recap: to improve learning and memory, we need to encode information in conjunction with excellent cues that will bring back the remembered events when we need them. But how do we do this? Keep in mind the two critical principles we have discussed: to maximize retrieval, we should construct *meaningful* cues that remind us of the original experience, and those cues should be *distinctive* and *not associated with other memories*. These two conditions are critical in maximizing cue effectiveness (Nairne, 2002).

So, how can these principles be adapted for use in many situations? Let's go back to how we started the module, with Simon Reinhard's ability to memorize huge numbers of digits. Although it was not obvious, he applied these same general memory principles, but in a more deliberate way. In fact, all **mnemonic devices**, or memory aids/tricks, rely on these fundamental principles. In a typical case, the person learns a set of cues and then applies these cues to learn and remember information. Consider the set of 20 items below that are easy to learn and remember (Bower & Reitman, 1972).

1. is a gun. 11 is penny-one, hot dog bun.
2. is a shoe. 12 is penny-two, airplane glue.
3. is a tree. 13 is penny-three, bumble bee.
4. is a door. 14 is penny-four, grocery store.
5. is knives. 15 is penny-five, big beehive.
6. is sticks. 16 is penny-six, magic tricks.
7. is oven. 17 is penny-seven, go to heaven.
8. is plate. 18 is penny-eight, golden gate.
9. is wine. 19 is penny-nine, ball of twine.
10. is hen. 20 is penny-ten, ballpoint pen.

It would probably take you less than 10 minutes to learn this list and practice recalling it several times (remember to use retrieval practice!). If you were to do so, you would have a set of peg words on which you could "hang" memories. In fact, this mnemonic device is called the *peg word technique*. If you then needed to remember some discrete items—say a grocery list, or points you wanted to make in a speech—this method would let you do so in a very precise yet flexible way. Suppose you had to remember bread, peanut butter, bananas, lettuce, and so on. The way to use the method is to form a vivid image of what you want to remember and imagine it interacting with your peg words (as many as you need). For example, for these items, you might imagine a large gun (the first peg word) shooting a loaf of bread, then a jar of peanut butter inside a shoe, then large bunches of bananas hanging from a tree, then a door slamming on a head of lettuce with leaves flying everywhere. The idea is to provide good, distinctive cues (the weirder the better!) for the information you need to remember while you are learning it. If you do this, then retrieving it later is relatively easy. You know your cues perfectly (one is gun, etc.), so you simply go through your cue word list and "look" in your mind's eye at the image stored there (bread, in this case).

This peg word method may sound strange at first, but it works quite well, even with little



Example of a mneumonic system created by a student to study cranial nerves. [Image: Kelidimari, https://goo.gl/kiA1kP, CC BY-SA 3.0, https://goo.gl/SCkRfm]

training (Roediger, 1980). One word of warning, though, is that the items to be remembered need to be presented relatively slowly at first, until you have practice associating each with its cue word. People get faster with time. Another interesting aspect of this technique is that it's just as easy to recall the items in backwards order as forwards. This is because the peg words provide direct access to the memorized items, regardless of order.

How did Simon Reinhard remember those digits? Essentially he has a much more complex system based on these same principles. In his case, he uses "memory palaces" (elaborate scenes with discrete places) combined with huge sets of images for digits. For example, imagine mentally walking through the home where you grew up and identifying as many distinct areas and objects as possible. Simon has hundreds of such memory palaces that he uses. Next, for remembering digits, he has memorized a set of 10,000 images. Every four-digit number for him immediately brings forth a mental image. So, for example, 6187 might recall Michael Jackson. When Simon hears all the numbers coming at him, he places an image for every four digits into locations in his memory palace. He can do this at an incredibly rapid rate, faster than 4 digits per 4 seconds when they are flashed visually, as in the demonstration at the beginning of the module. As noted, his record is 240 digits, recalled in exact order. Simon also holds the world record in an event called "speed cards," which involves memorizing the precise order of a shuffled deck of cards. Simon was able to do this in 21.19 seconds! Again, he uses his memory palaces, and he encodes groups of cards as single images.

Many books exist on how to improve memory using mnemonic devices, but all involve forming distinctive encoding operations and then having an infallible set of memory cues. We should add that to develop and use these memory systems beyond the basic peg system outlined above takes a great amount of time and concentration. The World Memory Championships are held every year and the records keep improving. However, for most common purposes, just keep in mind that to remember well you need to encode information in a distinctive way and to have good cues for retrieval. You can adapt a system that will meet most any purpose.

Outside Resources

Book: Brown, P.C., Roediger, H. L. & McDaniel, M. A. (2014). Make it stick: The science of successful learning. Cambridge, MA: Harvard University Press.
https://www.amazon.com/Make-Stick-Science-Successful-Learning/dp/0674729013

Student Video 1: Eureka Foong's - The Misinformation Effect. This is a student-made video illustrating this phenomenon of altered memory. It was one of the winning entries in the 2014 Noba Student Video Award.
https://www.youtube.com/watch?v=iMPIWkFtd88

Student Video 2: Ang Rui Xia & Ong Jun Hao's - The Misinformation Effect. Another student-made video exploring the misinformation effect. Also an award winner from 2014.
https://www.youtube.com/watch?v=gsn9iKmOJLQ

Video: Flashbulb Memories
https://youtu.be/3kE1M-MfXxc

Video: Simon Reinhard breaking the world record in speedcards.
http://vimeo.com/12516465

Web: Retrieval Practice, a website with research, resources, and tips for both educators and learners around the memory-strengthening skill of retrieval practice.
http://www.retrievalpractice.org/

Discussion Questions

1. Mnemonists like Simon Reinhard develop mental “journeys,” which enable them to use the method of loci. Develop your own journey, which contains 20 places, in order, that you know well. One example might be: the front walkway to your parents’ apartment; their doorbell; the couch in their living room; etc. Be sure to use a set of places that you know well and that have a natural order to them (e.g., the walkway comes before the doorbell). Now you are more than halfway toward being able to memorize a set of 20 nouns, in order, rather quickly. As an optional second step, have a friend make a list of 20 such nouns and read them to you, slowly (e.g., one every 5 seconds). Use the method to attempt to remember the 20 items.

2. Recall a recent argument or misunderstanding you have had about memory (e.g., a debate over whether your girlfriend/boyfriend had agreed to something). In light of what you have just learned about memory, how do you think about it? Is it possible that the disagreement can be understood by one of you making a pragmatic inference?
3. Think about what you've learned in this module and about how you study for tests. On the basis of what you have learned, is there something you want to try that might help your study habits?

Vocabulary

Autobiographical memory
Memory for the events of one's life.

Consolidation
The process occurring after encoding that is believed to stabilize memory traces.

Cue overload principle
The principle stating that the more memories that are associated to a particular retrieval cue, the less effective the cue will be in prompting retrieval of any one memory.

Distinctiveness
The principle that unusual events (in a context of similar events) will be recalled and recognized better than uniform (nondistinctive) events.

Encoding
The initial experience of perceiving and learning events.

Encoding specificity principle
The hypothesis that a retrieval cue will be effective to the extent that information encoded from the cue overlaps or matches information in the engram or memory trace.

Engrams
A term indicating the change in the nervous system representing an event; also, memory trace.

Episodic memory
Memory for events in a particular time and place.

Flashbulb memory
Vivid personal memories of receiving the news of some momentous (and usually emotional) event.

Memory traces
A term indicating the change in the nervous system representing an event.

Misinformation effect
When erroneous information occurring after an event is remembered as having been part of

the original event.

Mnemonic devices
A strategy for remembering large amounts of information, usually involving imaging events occurring on a journey or with some other set of memorized cues.

Recoding
The ubiquitous process during learning of taking information in one form and converting it to another form, usually one more easily remembered.

Retrieval
The process of accessing stored information.

Retroactive interference
The phenomenon whereby events that occur after some particular event of interest will usually cause forgetting of the original event.

Semantic memory
The more or less permanent store of knowledge that people have.

Storage
The stage in the learning/memory process that bridges encoding and retrieval; the persistence of memory over time.

References

Anderson, M. C., Bjork, R., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 20*, 1063–1087.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge: Cambridge University Press.

Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior, 11*, 478–485.

Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition, 5(6)*, 673–678.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*, 73–99.

Chan, J.C.K. & McDermott, K.B. (2006). Remembering pragmatic inferences. *Applied Cognitive Psychology, 20*, 633-639.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*,66 (3), 325-331

Hunt, R. (2003). Two contributions of distinctive processing to accurate memory. *Journal of Memory and Language, 48*, 811–825.

Hunt, R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32*, 421-445.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361–366.

McDermott, K. B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false recall. *Memory & Cognition, 34*, 261–267.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39(4)*, 352.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior, 2*, 1–21.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Muter, P. (1984). Recognition and recall of words with a single meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 198–202.

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10*, 389–395.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*, 838–848.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*, 437–447.

Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 558.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology-Learning Memory and Cognition, 21*, 803–814.

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition, 27*, 494–500.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science, 14*, 455–461.

Tulving, E. (2007). Are there 256 different kinds of memory? In J.S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 39–52). New York: Psychology Press.

Tulving, E. (1991). Interview. *Journal of Cognitive Neuroscience, 3*, 89–94

Tulving, E., & Bower, G. H. (1975). The logic of memory representations. *The psychology of learning and motivation, 8*, 265-301.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5*, 381–391.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Watkins, M. J. (1975). Inhibition in recall with extralist "cues." *Journal of Verbal Learning and Verbal Behavior, 14*, 294–303.

Watkins, M. J., & Tulving, E. (1975). Episodic memory: When recognition fails. *Journal of Experimental Psychology: General, 104*, 5–29.

21
Conditioning and Learning

Mark E. Bouton

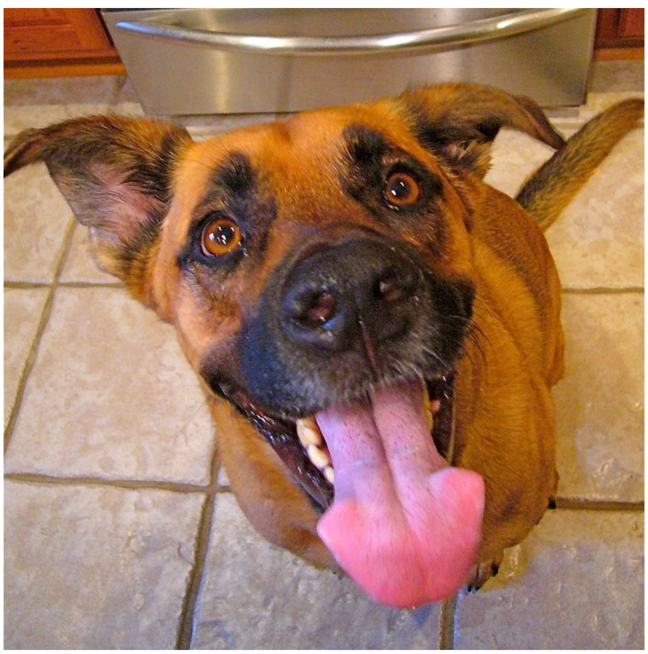
Basic principles of learning are always operating and always influencing human behavior. This module discusses the two most fundamental forms of learning -- classical (Pavlovian) and instrumental (operant) conditioning. Through them, we respectively learn to associate 1) stimuli in the environment, or 2) our own behaviors, with significant events, such as rewards and punishments. The two types of learning have been intensively studied because they have powerful effects on behavior, and because they provide methods that allow scientists to analyze learning processes rigorously. This module describes some of the most important things you need to know about classical and instrumental conditioning, and it illustrates some of the many ways they help us understand normal and disordered behavior in humans. The module concludes by introducing the concept of observational learning, which is a form of learning that is largely distinct from classical and operant conditioning.

Learning Objectives

- Distinguish between classical (Pavlovian) conditioning and instrumental (operant) conditioning.
- Understand some important facts about each that tell us how they work.
- Understand how they work separately and together to influence human behavior in the world outside the laboratory.
- Students will be able to list the four aspects of observational learning according to Social Learning Theory.

Two Types of Conditioning

Although Ivan Pavlov won a Nobel Prize for studying digestion, he is much more famous for something else: working with a dog, a bell, and a bowl of saliva. Many people are familiar with the classic study of "Pavlov's dog," but rarely do they understand the significance of its discovery. In fact, Pavlov's work helps explain why some people get anxious just looking at a crowded bus, why the sound of a morning alarm is so hated, and even why we swear off certain foods we've only tried once. Classical (or Pavlovian) conditioning is one of the fundamental ways we learn about the world around us. But it is far more than just a theory of learning; it is also arguably a theory of identity. For, once you understand classical conditioning, you'll recognize that your favorite music, clothes, even political candidate, might all be a result of the same process that makes a dog drool at the sound of bell.



The Pavlov in All of Us: Does your dog learn to beg for food because you reinforce her by feeding her from the table? [Image: David Mease, https://goo.gl/R9cQV7, CC BY-NC 2.0, https://goo.gl/FIlc2e]

Around the turn of the 20th century, scientists who were interested in understanding the behavior of animals and humans began to appreciate the importance of two very basic forms of learning. One, which was first studied by the Russian physiologist Ivan Pavlov, is known as **classical**, or **Pavlovian conditioning**. In his famous experiment, Pavlov rang a bell and then gave a dog some food. After repeating this pairing multiple times, the dog eventually treated the bell as a signal for food, and began salivating in anticipation of the treat. This kind of result has been reproduced in the lab using a wide range of signals (e.g., tones, light, tastes, settings) paired with many different events besides food (e.g., drugs, shocks, illness; see below).

We now believe that this same learning process is engaged, for example, when humans associate a drug they've taken with the environment in which they've taken it; when they associate a stimulus (e.g., a symbol for vacation, like a big beach towel) with an emotional event (like a burst of happiness); and when they associate the flavor of a food with getting food poisoning. Although classical conditioning may seem "old" or "too simple" a theory, it is still widely studied today for at least two reasons: First, it is a straightforward test of associative learning that can be used to study other, more complex behaviors. Second, because classical conditioning is always occurring in our lives, its effects on behavior have important implications for understanding normal and disordered behavior in humans.

In a general way, classical conditioning occurs whenever neutral stimuli are associated with psychologically significant events. With food poisoning, for example, although having fish for dinner may not normally be something to be concerned about (i.e., a "neutral stimuli"), if it causes you to get sick, you will now likely associate that neutral stimuli (the fish) with the psychologically significant event of getting sick. These paired events are often described using terms that can be applied to any situation.

The dog food in Pavlov's experiment is called the **unconditioned stimulus (US)** because it elicits an **unconditioned response (UR)**. That is, without any kind of "training" or "teaching," the stimulus produces a natural or instinctual reaction. In Pavlov's case, the food (US) automatically makes the dog drool (UR). Other examples of unconditioned stimuli include loud noises (US) that startle us (UR), or a hot shower (US) that produces pleasure (UR).

On the other hand, a conditioned stimulus produces a conditioned response. A **conditioned stimulus (CS)** is a signal that has no importance to the organism until it is paired with something that does have importance. For example, in Pavlov's experiment, the bell is the conditioned stimulus. Before the dog has learned to associate the bell (CS) with the presence of food (US), hearing the bell means nothing to the dog. However, after multiple pairings of the bell with the presentation of food, the dog starts to drool at the sound of the bell. This drooling in response to the bell is the **conditioned response (CR)**. Although it can be confusing, the conditioned response is almost always the same as the unconditioned response. However, it is called the conditioned response because it is conditional on (or, depends on) being paired with the conditioned stimulus (e.g., the bell). To help make this clearer, consider becoming really hungry when you see the logo for a fast food restaurant. There's a good chance you'll start salivating. Although it is the actual eating of the food (US) that normally produces the salivation (UR), simply seeing the restaurant's logo (CS) can trigger the same reaction (CR).

Another example you are probably very familiar with involves your alarm clock. If you're like most people, waking up early usually makes you unhappy. In this case, waking up early (US) produces a natural sensation of grumpiness (UR). Rather than waking up early on your own, though, you likely have an alarm clock that plays a tone to wake you. Before setting your alarm to that particular tone, let's imagine you had neutral feelings about it (i.e., the tone had no prior meaning for you). However, now that you use it to wake up every morning, you psychologically "pair" that tone (CS) with your feelings of grumpiness in the morning (UR).

After enough pairings, this tone (CS) will automatically produce your natural response of grumpiness (CR). Thus, this linkage between the unconditioned stimulus (US; waking up early) and the conditioned stimulus (CS; the tone) is so strong that the unconditioned response (UR; being grumpy) will become a conditioned response (CR; e.g., hearing the tone at any point in the day—whether waking up or walking down the street—will make you grumpy). Modern studies of classical conditioning use a very wide range of CSs and USs and measure a wide range of conditioned responses.

Receiving a reward can condition you toward certain behaviors. For example, when you were a child, your mother may have offered you this deal: "Don't make a fuss when we're in the supermarket and you'll get a treat on the way out." [Image: Oliver Hammond, https://goo.gl/xFKiZL, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Although classical conditioning is a powerful explanation for how we learn many different things, there is a second form of conditioning that also helps explain how we learn. First studied by Edward Thorndike, and later extended by B. F. Skinner, this second type of conditioning is known as **instrumental** or **operant conditioning**. Operant conditioning occurs when a *behavior* (as opposed to a stimulus) is associated with the occurrence of a significant event. In the best-known example, a rat in a laboratory learns to press a lever in a cage (called a "Skinner box") to receive food. Because the rat has no "natural" association between pressing a lever and getting food, the rat has to learn this connection. At first, the rat may simply explore its cage, climbing on top of things, burrowing under things, in search of food. Eventually while poking around its cage, the rat accidentally presses the lever, and a food pellet drops in. This voluntary behavior is called an **operant** behavior, because it "operates" on the environment (i.e., it is an action that the animal itself makes).

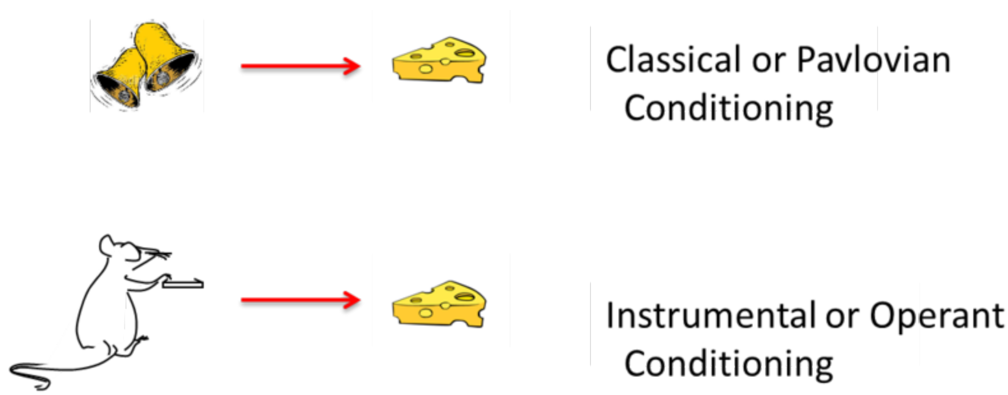
Now, once the rat recognizes that it receives a piece of food every time it presses the lever, the behavior of lever-pressing becomes reinforced. That is, the food pellets serve as **reinforcer**s because they strengthen the rat's desire to engage with the environment in this particular manner. In a parallel example, imagine that you're playing a street-racing video game. As you drive through one city course multiple times, you try a number of different streets to get to the finish line. On one of these trials, you discover a shortcut that dramatically improves your overall time. You have learned this new path through operant conditioning.

That is, by engaging with your environment (operant responses), you performed a sequence of behaviors that that was positively reinforced (i.e., you found the shortest distance to the finish line). And now that you've learned how to drive this course, you will perform that same sequence of driving behaviors (just as the rat presses on the lever) to receive your reward of a faster finish.

Operant conditioning research studies how the effects of a behavior influence the probability that it will occur again. For example, the effects of the rat's lever-pressing behavior (i.e., receiving a food pellet) influences the probability that it will keep pressing the lever. For, according to Thorndike's law of effect, when a behavior has a positive (satisfying) effect or consequence, it is likely to be repeated in the future. However, when a behavior has a negative (painful/annoying) consequence, it is less likely to be repeated in the future. Effects that increase behaviors are referred to as reinforcers, and effects that decrease them are referred to as punishers.

An everyday example that helps to illustrate operant conditioning is striving for a good grade in class—which could be considered a reward for students (i.e., it produces a positive emotional response). In order to get that reward (similar to the rat learning to press the lever), the student needs to modify his/her behavior. For example, the student may learn that speaking up in class gets him/her participation points (a reinforcer), so the student speaks up repeatedly. However, the student also learns that s/he shouldn't speak up about just anything; talking about topics unrelated to school actually costs points. Therefore, through the student's freely chosen behaviors, s/he learns which behaviors are reinforced and which are punished.

An important distinction of operant conditioning is that it provides a method for studying how consequences influence "voluntary" behavior. The rat's decision to press the lever is voluntary, in the sense that the rat is free to make and repeat that response whenever it wants. Classical



[Image courtesy of Bernard W. Balleine]

conditioning, on the other hand, is just the opposite—depending instead on "involuntary" behavior (e.g., the dog doesn't choose to drool; it just does). So, whereas the rat must actively participate and perform some kind of behavior to attain its reward, the dog in Pavlov's experiment is a passive participant. One of the lessons of operant conditioning research, then, is that voluntary behavior is strongly influenced by its consequences.

The illustration above summarizes the basic elements of classical and instrumental conditioning. The two types of learning differ in many ways. However, modern thinkers often emphasize the fact that they differ—as illustrated here—in *what* is learned. In classical conditioning, the animal behaves as if it has learned to associate a *stimulus* with a significant event. In operant conditioning, the animal behaves as if it has learned to associate a *behavior* with a significant event. Another difference is that the response in the classical situation (e.g., salivation) is *elicited* by a stimulus that comes before it, whereas the response in the operant case is not elicited by any particular stimulus. Instead, operant responses are said to be *emitted*. The word "emitted" further conveys the idea that operant behaviors are essentially voluntary in nature.

Understanding classical and operant conditioning provides psychologists with many tools for understanding learning and behavior in the world outside the lab. This is in part because the two types of learning occur continuously throughout our lives. It has been said that "much like the laws of gravity, the laws of learning are always in effect" (Spreat & Spreat, 1982).

Useful Things to Know about Classical Conditioning

Classical Conditioning Has Many Effects on Behavior

A classical CS (e.g., the bell) does not merely elicit a simple, unitary reflex. Pavlov emphasized salivation because that was the only response he measured. But his bell almost certainly elicited a whole *system* of responses that functioned to get the organism ready for the upcoming US (food) (see Timberlake, 2001). For example, in addition to salivation, CSs (such as the bell) that signal that food is near also elicit the secretion of gastric acid, pancreatic enzymes, and insulin (which gets blood glucose into cells). All of these responses prepare the body for digestion. Additionally, the CS elicits approach behavior and a state of excitement. And presenting a CS for food can also cause animals whose stomachs are full to eat more food if it is available. In fact, food CSs are so prevalent in modern society, humans are likewise inclined to eat or feel hungry in response to cues associated with food, such as the sound of a bag of potato chips opening, the sight of a well-known logo (e.g., Coca-Cola), or the feel of the couch in front of the television.

Classical conditioning is also involved in other aspects of eating. Flavors associated with certain nutrients (such as sugar or fat) can become preferred without arousing any awareness of the pairing. For example, protein is a US that your body automatically craves more of once you start to consume it (UR): since proteins are highly concentrated in meat, the flavor of meat becomes a CS (or cue, that proteins are on the way), which perpetuates the cycle of craving for yet more meat (this automatic bodily reaction now a CR).

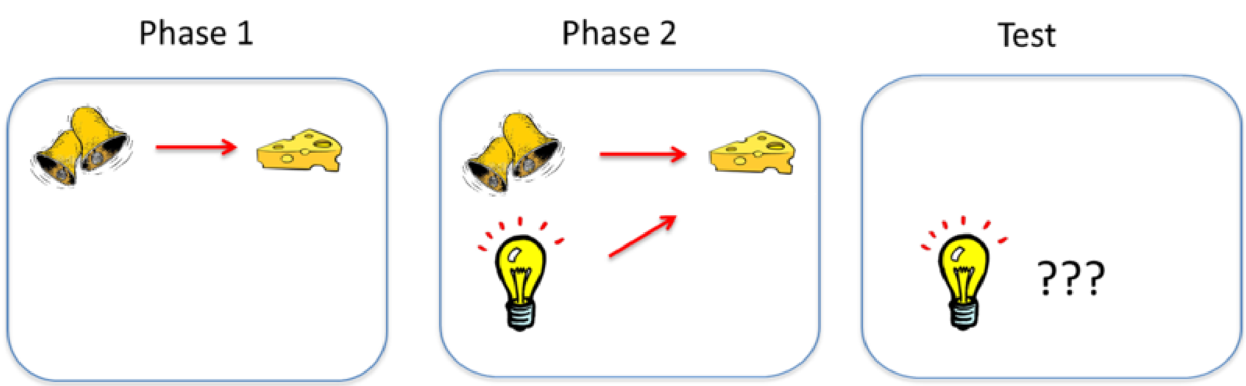
In a similar way, flavors associated with stomach pain or illness become avoided and *dis*liked. For example, a person who gets sick after drinking too much tequila may acquire a profound dislike of the taste and odor of tequila—a phenomenon called **taste aversion conditioning**. The fact that flavors are often associated with so many consequences of eating is important for animals (including rats and humans) that are frequently exposed to new foods. And it is clinically relevant. For example, drugs used in chemotherapy often make cancer patients sick. As a consequence, patients often acquire aversions to foods eaten just before treatment, or even aversions to such things as the waiting room of the chemotherapy clinic itself (see Bernstein, 1991; Scalera & Bavieri, 2009).

Classical conditioning occurs with a variety of significant events. If an experimenter sounds a tone just before applying a mild shock to a rat's feet, the tone will elicit fear or anxiety after one or two pairings. Similar **fear conditioning** plays a role in creating many anxiety disorders in humans, such as phobias and panic disorders, where people associate cues (such as closed spaces, or a shopping mall) with panic or other emotional trauma (see Mineka & Zinbarg, 2006). Here, rather than a physical response (like drooling), the CS triggers an emotion.

Another interesting effect of classical conditioning can occur when we ingest drugs. That is, when a drug is taken, it can be associated with the cues that are present at the same time (e. g., rooms, odors, drug paraphernalia). In this regard, if someone associates a particular smell with the sensation induced by the drug, whenever that person smells the same odor afterward, it may cue responses (physical and/or emotional) related to taking the drug itself. But drug cues have an even more interesting property: They elicit responses that often "compensate" for the upcoming effect of the drug (see Siegel, 1989). For example, morphine itself suppresses pain; however, if someone is used to taking morphine, a cue that signals the "drug is coming soon" can actually make the person more sensitive to pain. Because the person knows a pain suppressant will soon be administered, the body becomes more sensitive, anticipating that "the drug will soon take care of it." Remarkably, such **conditioned compensatory responses** in turn decrease the impact of the drug on the body—because the body has become more sensitive to pain.

This conditioned compensatory response has many implications. For instance, a drug user will be most "tolerant" to the drug in the presence of cues that have been associated with it (because such cues elicit compensatory responses). As a result, overdose is usually not due to an increase in dosage, but to taking the drug in a new place without the familiar cues—which would have otherwise allowed the user to tolerate the drug (see Siegel, Hinson, Krank, & McCully, 1982). Conditioned compensatory responses (which include heightened pain sensitivity and decreased body temperature, among others) might also cause discomfort, thus motivating the drug user to continue usage of the drug to reduce them. This is one of several ways classical conditioning might be a factor in drug addiction and dependence.

A final effect of classical cues is that they motivate ongoing operant behavior (see Balleine, 2005). For example, if a rat has learned via operant conditioning that pressing a lever will give it a drug, in the presence of cues that signal the "drug is coming soon" (like the sound of the lever squeaking), the rat will work harder to press the lever than if those cues weren't present (i.e., there is no squeaking lever sound). Similarly, in the presence of food-associated cues (e. g., smells), a rat (or an overeater) will work harder for food. And finally, even in the presence of negative cues (like something that signals fear), a rat, a human, or any other organism will work harder to avoid those situations that might lead to trauma. Classical CSs thus have many effects that can contribute to significant behavioral phenomena.



[Image courtesy of Bernard W. Balleine]

The Learning Process

As mentioned earlier, classical conditioning provides a method for studying basic learning processes. Somewhat counterintuitively, though, studies show that pairing a CS and a US together is not sufficient for an association to be learned between them. Consider an effect called blocking (see Kamin, 1969). In this effect, an animal first learns to associate one CS—call it stimulus A—with a US. In the illustration above, the sound of a bell (stimulus A) is paired with the presentation of food. Once this association is learned, in a second phase, a second stimulus—stimulus B—is presented alongside stimulus A, such that the two stimuli are paired with the US together. In the illustration, a light is added and turned on at the same time the bell is rung. However, because the animal has already learned the association between stimulus A (the bell) and the food, the animal doesn't learn an association between stimulus B (the light) and the food. That is, the conditioned response only occurs during the presentation of stimulus A, because the earlier conditioning of A "blocks" the conditioning of B when B is added to A. The reason? Stimulus A already predicts the US, so the US is not surprising when it occurs with Stimulus B.

Learning depends on such a surprise, or a discrepancy between what occurs on a conditioning trial and what is already predicted by cues that are present on the trial. To learn something through classical conditioning, there must first be some **prediction error**, or the chance that a conditioned stimulus won't lead to the expected outcome. With the example of the bell and the light, because the bell always leads to the reward of food, there's no "prediction error" that the addition of the light helps to correct. However, if the researcher suddenly requires that the bell and the light both occur in order to receive the food, the bell alone will produce a prediction error that the animal has to learn.

Blocking and other related effects indicate that the learning process tends to take in the most valid predictors of significant events and ignore the less useful ones. This is common in the real world. For example, imagine that your supermarket puts big star-shaped stickers on products that are on sale. Quickly, you learn that items with the big star-shaped stickers are cheaper. However, imagine you go into a similar supermarket that not only uses these stickers, but also uses bright orange price tags to denote a discount. Because of blocking (i.e., you already know that the star-shaped stickers indicate a discount), you don't have to learn the color system, too. The star-shaped stickers tell you everything you need to know (i.e. there's no prediction error for the discount), and thus the color system is irrelevant.

Classical conditioning is strongest if the CS and US are intense or salient. It is also best if the CS and US are relatively new and the organism hasn't been frequently exposed to them before. And it is especially strong if the organism's biology has prepared it to associate a particular CS and US. For example, rats and humans are naturally inclined to associate an illness with a flavor, rather than with a light or tone. Because foods are most commonly experienced by taste, if there is a particular food that makes us ill, associating the flavor (rather than the appearance—which may be similar to other foods) with the illness will more greatly ensure we avoid that food in the future, and thus avoid getting sick. This sorting tendency, which is set up by evolution, is called **preparedness**.

There are many factors that affect the strength of classical conditioning, and these have been the subject of much research and theory (see Rescorla & Wagner, 1972; Pearce & Bouton, 2001). Behavioral neuroscientists have also used classical conditioning to investigate many of the basic brain processes that are involved in learning (see Fanselow & Poulos, 2005; Thompson & Steinmetz, 2009).

Erasing Classical Learning

After conditioning, the response to the CS can be eliminated if the CS is presented repeatedly without the US. This effect is called **extinction**, and the response is said to become "extinguished." For example, if Pavlov kept ringing the bell but never gave the dog any food afterward, eventually the dog's CR (drooling) would no longer happen when it heard the CS (the bell), because the bell would no longer be a predictor of food. Extinction is important for many reasons. For one thing, it is the basis for many therapies that clinical psychologists use to eliminate maladaptive and unwanted behaviors. Take the example of a person who has a debilitating fear of spiders: one approach might include systematic exposure to spiders. Whereas, initially the person has a CR (e.g., extreme fear) every time s/he sees the CS (e.g., the spider), after repeatedly being shown pictures of spiders in neutral conditions, pretty soon the CS no longer predicts the CR (i.e., the person doesn't have the fear reaction when seeing spiders, having learned that spiders no longer serve as a "cue" for that fear). Here, repeated exposure to spiders without an aversive consequence causes extinction.

Psychologists must accept one important fact about extinction, however: it does not necessarily destroy the original learning (see Bouton, 2004). For example, imagine you strongly associate the smell of chalkboards with the agony of middle school detention. Now imagine that, after years of encountering chalkboards, the smell of them no longer recalls the agony of detention (an example of extinction). However, one day, after entering a new building for the first time, you suddenly catch a whiff of a chalkboard and WHAM!, the agony of detention returns. This is called **spontaneous recovery**: following a lapse in exposure to the CS after extinction has occurred, sometimes re-exposure to the CS (e.g., the smell of chalkboards) can evoke the CR again (e.g., the agony of detention).

Another related phenomenon is the **renewal effect**: After extinction, if the CS is tested in a new **context**, such as a different room or location, the CR can also return. In the chalkboard example, the action of entering a new building—where you don't expect to smell chalkboards—suddenly renews the sensations associated with detention. These effects have been interpreted to suggest that extinction *inhibits* rather than erases the learned behavior, and this inhibition is mainly expressed in the context in which it is learned (see "context" in the

Key Vocabulary section below).

This does not mean that extinction is a bad treatment for behavior disorders. Instead, clinicians can increase its effectiveness by using basic research on learning to help defeat these relapse effects (see Craske et al., 2008). For example, conducting extinction therapies in contexts where patients might be most vulnerable to relapsing (e.g., at work), might be a good strategy for enhancing the therapy's success.

Useful Things to Know about Instrumental Conditioning

Most of the things that affect the strength of classical conditioning also affect the strength of instrumental learning—whereby we learn to associate our actions with their outcomes. As noted earlier, the "bigger" the reinforcer (or punisher), the stronger the learning. And, if an instrumental behavior is no longer reinforced, it will also be extinguished. Most of the rules of associative learning that apply to classical conditioning also apply to instrumental learning, but other facts about instrumental learning are also worth knowing.

Instrumental Responses Come Under Stimulus Control

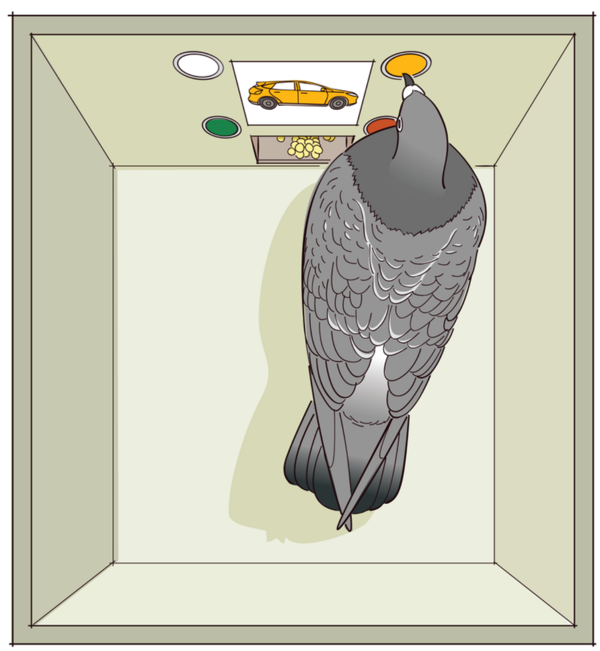
As you know, the classic operant response in the laboratory is lever-pressing in rats, reinforced by food. However, things can be arranged so that lever-pressing only produces pellets when a particular stimulus is present. For example, lever-pressing can be reinforced only when a light in the Skinner box is turned on; when the light is off, no food is released from lever-pressing. The rat soon learns to discriminate between the light-on and light-off conditions, and presses the lever only in the presence of the light (responses in light-off are extinguished). In everyday life, think about waiting in the turn lane at a traffic light. Although you know that green means go, only when you have the green *arrow* do you turn. In this regard, the operant behavior is now said to be under **stimulus control**. And, as is the case with the traffic light, in the real world, stimulus control is probably the rule.

The stimulus controlling the operant response is called a **discriminative stimulus**. It can be associated directly with the response, or the reinforcer (see below). However, it usually does not elicit the response the way a classical CS does. Instead, it is said to "set the occasion for" the operant response. For example, a canvas put in front of an artist does not elicit painting behavior or compel her to paint. It allows, or sets the occasion for, painting to occur.

Stimulus-control techniques are widely used in the laboratory to study perception and other psychological processes in animals. For example, the rat would not be able to respond appropriately to light-on and light-off conditions if it could not see the light. Following this logic, experiments using stimulus-control methods have tested how well animals see colors, hear ultrasounds, and detect magnetic fields. That is, researchers pair these discriminative stimuli with those they know the animals already understand (such as pressing the lever). In this way, the researchers can test if the animals can learn to press the lever only when an ultrasound is played, for example.

These methods can also be used to study "higher" cognitive processes. For example, pigeons can learn to peck at different buttons in a Skinner box when pictures of flowers, cars, chairs, or people are shown on a miniature TV screen (see Wasserman, 1995). Pecking button 1 (and no other) is reinforced in the presence of a flower image, button 2 in the presence of a chair image, and so on. Pigeons can learn the discrimination readily, and, under the right conditions, will even peck the correct buttons associated with pictures of *new* flowers, cars, chairs, and people they have never seen before. The birds have learned to categorize the sets of stimuli. Stimulus-control methods can be used to study how such categorization is learned.

Operant Conditioning Involves Choice

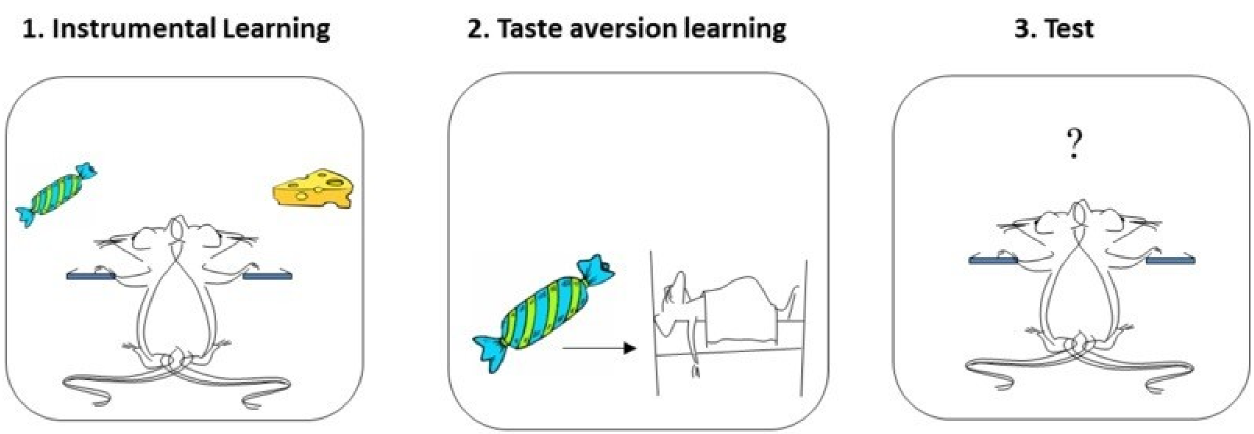


Another thing to know about operant conditioning is that the response always requires choosing one behavior over others. The student who goes to the bar on Thursday night chooses to drink instead of staying at home and studying. The rat chooses to press the lever instead of sleeping or scratching its ear in the back of the box. The alternative behaviors are each associated with their own reinforcers. And the tendency to perform a particular action depends on both the reinforcers earned for it and the reinforcers earned for its alternatives.

To investigate this idea, choice has been studied in the Skinner box by making two levers available for the rat (or two buttons available for the pigeon), each of which has its own reinforcement or payoff rate. A thorough study of choice in situations like this has led to a rule called the quantitative law of effect (see Herrnstein, 1970), which can be understood without going into quantitative detail: The law acknowledges the fact that the effects of reinforcing one behavior depend crucially on how much reinforcement is earned for the behavior's alternatives. For example, if a pigeon learns that pecking one light will reward two food pellets, whereas the other light only rewards one, the pigeon will only peck the first light. However, what happens if the first light is more strenuous to reach than the second one? Will the cost of energy outweigh the bonus of food? Or will the extra food be worth the work? In general, a given reinforcer will be less reinforcing if there are many alternative reinforcers in the environment. For this reason, alcohol, sex, or drugs may be less powerful reinforcers if the person's environment is full of other sources of reinforcement, such as achievement at work or love from family members.

Cognition in Instrumental Learning

Modern research also indicates that reinforcers do more than merely strengthen or "stamp in" the behaviors they are a consequence of, as was Thorndike's original view. Instead, animals learn about the specific consequences of each behavior, and will perform a behavior depending on how much they currently want—or "value"—its consequence.



[Image courtesy of Bernard W. Balleine]

This idea is best illustrated by a phenomenon called the reinforcer devaluation effect (see Colwill & Rescorla, 1986). A rat is first trained to perform two instrumental actions (e.g., pressing a lever on the left, and on the right), each paired with a different reinforcer (e.g., a sweet sucrose solution, and a food pellet). At the end of this training, the rat tends to press both levers, alternating between the sucrose solution and the food pellet. In a second phase, one of the reinforcers (e.g., the sucrose) is then separately paired with illness. This conditions a taste aversion to the sucrose. In a final test, the rat is returned to the Skinner box and allowed to press either lever freely. No reinforcers are presented during this test (i.e., no sucrose or food comes from pressing the levers), so behavior during testing can only result from the rat's memory of what it has learned earlier. Importantly here, the rat chooses *not* to perform the response that once produced the reinforcer that it now has an aversion to (e.g., it won't press the sucrose lever). This means that the rat has learned and remembered the reinforcer associated with each response, and can combine that knowledge with the knowledge that the reinforcer is now "bad." Reinforcers do not merely stamp in responses; the animal learns much more than that. The behavior is said to be "**goal-directed**" (see Dickinson & Balleine, 1994), because it is influenced by the current value of its associated goal (i.e., how much the rat wants/doesn't want the reinforcer).

Things can get more complicated, however, if the rat performs the instrumental actions frequently and repeatedly. That is, if the rat has spent many months learning the value of pressing each of the levers, the act of pressing them becomes automatic and routine. And here, this once goal-directed action (i.e., the rat pressing the lever for the goal of getting sucrose/food) can become a **habit**. Thus, if a rat spends many months performing the lever-pressing behavior (turning such behavior into a habit), even when sucrose is again paired with illness, the rat will continue to press that lever (see Holland, 2004). After all the practice, the instrumental response (pressing the lever) is no longer sensitive to reinforcer devaluation. The rat continues to respond automatically, regardless of the fact that the sucrose from this lever makes it sick.

Habits are very common in human experience, and can be useful. You do not need to relearn each day how to make your coffee in the morning or how to brush your teeth. Instrumental behaviors can eventually become habitual, letting us get the job done while being free to think about other things.

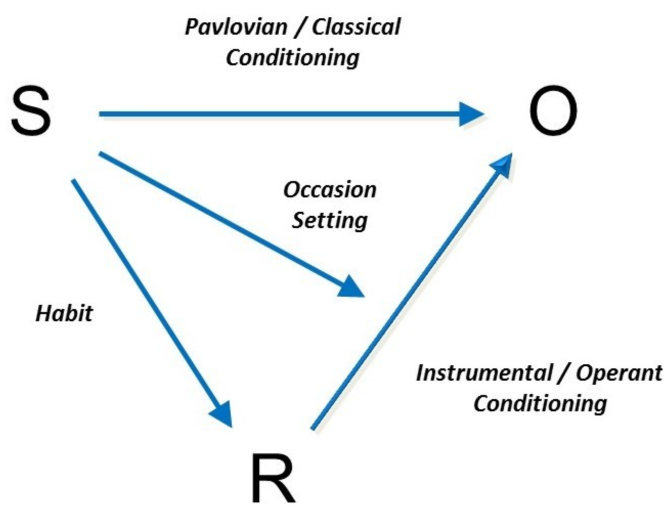
Putting Classical and Instrumental Conditioning Together

Classical and operant conditioning are usually studied separately. But outside of the laboratory they almost always occur at the same time. For example, a person who is reinforced for drinking alcohol or eating excessively learns these behaviors in the presence of certain stimuli—a pub, a set of friends, a restaurant, or possibly the couch in front of the TV. These stimuli are also available for association with the reinforcer. In this way, classical and operant conditioning are always intertwined.

The figure below summarizes this idea, and helps review what we have discussed in this module. Generally speaking, any reinforced or punished operant response (R) is paired with an outcome (O) in the presence of some stimulus or set of stimuli (S).

The figure illustrates the types of associations that can be learned in this very general scenario. For one thing, the organism will learn to associate the response *and* the outcome (R – O). This is instrumental conditioning. The learning process here is probably similar to classical conditioning, with all its emphasis on surprise and prediction error. And, as we discussed while considering the reinforcer devaluation effect, once R – O is learned, the organism will be ready to perform the response if the outcome is desired or valued. The value of the reinforcer can also be influenced by other reinforcers earned for other behaviors in the situation. These factors are at the heart of instrumental learning.

Second, the organism can also learn to associate the stimulus with the reinforcing outcome (S – O). This is the classical conditioning component, and as we have seen, it can have many consequences on behavior. For one thing, the stimulus will come to evoke a system of responses that help the organism prepare for the reinforcer (not shown in the figure): The drinker may undergo changes in body temperature; the eater may salivate and have an increase in insulin secretion. In addition, the stimulus will evoke approach (if the outcome is positive) or retreat (if the outcome is negative). Presenting the stimulus will also prompt the instrumental response.



The third association in the diagram is the one between the stimulus and the response (S – R). As discussed earlier, after a lot of practice, the stimulus may begin to elicit the response directly. This is habit learning, whereby the response occurs relatively automatically, without much mental processing of the relation between the action and the outcome and the outcome's current value.

The final link in the figure is between the stimulus and the response-outcome association [S – (R – O)]. More than just entering into a simple association with the R or the O, the stimulus can signal that the R – O relationship is now in effect. This is what we mean when we say that the stimulus can "set the occasion" for the operant response: It sets the occasion for the response-reinforcer relationship. Through this mechanism, the painter might begin to paint when given the right tools and the opportunity enabled by the canvas. The canvas theoretically signals that the behavior of painting will now be reinforced by positive consequences.

The figure provides a framework that you can use to understand almost any learned behavior you observe in yourself, your family, or your friends. If you would like to understand it more deeply, consider taking a course on learning in the future, which will give you a fuller appreciation of how classical learning, instrumental learning, habit learning, and occasion setting actually work and interact.

Observational Learning

Not all forms of learning are accounted for entirely by classical and operant conditioning. Imagine a child walking up to a group of children playing a game on the playground. The game looks fun, but it is new and unfamiliar. Rather than joining the game immediately, the child opts to sit back and watch the other children play a round or two. Observing the others, the child takes note of the ways in which they behave while playing the game. By watching the behavior of the other kids, the child can figure out the rules of the game and even some strategies for doing well at the game. This is called **observational learning**.

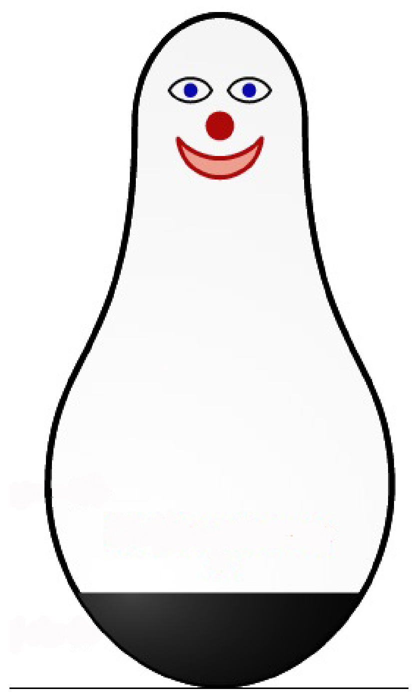
Children observing a social model (an experienced chess player) to learn the rules and strategies of the game of chess. [Image: David R. Tribble, https://goo.gl/nWsgxI, CC BY-SA 3.0, https://goo.gl/uhHola]

Observational learning is a component of Albert Bandura's **Social Learning Theory** (Bandura, 1977), which posits that individuals can learn novel responses via observation of key others' behaviors. Observational learning does not necessarily require reinforcement, but instead hinges on the presence of others, referred to as **social models**. Social models are typically of higher status or authority compared to the observer, examples of which include parents, teachers, and police officers. In the example above, the children who already know how to play the game could be thought of as being authorities—and are therefore social models—even though they are the same age as the observer. By observing how the social models behave, an individual is able to learn how to act in a certain situation. Other examples of observational learning might include a child learning to place her napkin in her lap by watching her parents at the dinner table, or a customer learning where to find the ketchup and mustard after observing other customers at a hot dog stand.

Bandura theorizes that the observational learning process consists of four parts. The first is *attention*—as, quite simply, one must pay attention to what s/he is observing in order to learn. The second part is *retention*: to learn one must be able to retain the behavior s/he is observing in memory.The third part of observational learning, *initiation*, acknowledges that the learner must be able to execute (or initiate) the learned behavior. Lastly, the observer must possess the *motivation* to engage in observational learning. In our vignette, the child must want to learn how to play the game in order to properly engage in observational learning.

Researchers have conducted countless experiments designed to explore observational learning, the most famous of which is Albert Bandura's "Bobo doll experiment."

In this experiment (Bandura, Ross & Ross 1961), Bandura had children individually observe an adult social model interact with a clown doll ("Bobo"). For one group of children, the adult interacted aggressively with Bobo: punching it, kicking it, throwing it, and even hitting it in the face with a toy mallet. Another group of children watched the adult interact with other toys, displaying no aggression toward Bobo. In both instances the adult left and the children were allowed to interact with Bobo on their own. Bandura found that children exposed to the aggressive social model were significantly more likely to behave aggressively toward Bobo, hitting and kicking him, compared to those exposed to the non-aggressive model. The researchers concluded that the children in the aggressive group used their observations of the adult social model's behavior to determine that aggressive behavior toward Bobo was acceptable.



Bobo [Image: © Sémhur / Wikimedia Commons / CC-BY-SA-3.0 (or Free Art License), https://goo.gl/uhHola]

While reinforcement was not required to elicit the children's behavior in Bandura's first experiment, it is important to acknowledge that consequences do play a role within observational learning. A future adaptation of this study (Bandura, Ross, & Ross, 1963) demonstrated that children in the aggression group showed less aggressive behavior if they witnessed the adult model receive punishment for aggressing against Bobo. Bandura referred to this process as **vicarious reinforcement**, as the children did not experience the reinforcement or punishment directly, yet were still influenced by observing it.

Conclusion

We have covered three primary explanations for how we learn to behave and interact with the world around us. Considering your own experiences, how well do these theories apply to you? Maybe when reflecting on your personal sense of fashion, you realize that you tend to select clothes others have complimented you on (operant conditioning). Or maybe, thinking back on a new restaurant you tried recently, you realize you chose it because its commercials play happy music (classical conditioning). Or maybe you are now always on time with your assignments, because you saw how others were punished when they were late (observational learning). Regardless of the activity, behavior, or response, there's a good chance your "decision" to do it can be explained based on one of the theories presented in this module.

Outside Resources

Article: Rescorla, R. A. (1988). Pavlovian conditioning: It’s not what you think it is. American Psychologist, 43, 151–160.

Book: Bouton, M. E. (2007). Learning and behavior: A contemporary synthesis. Sunderland, MA: Sinauer Associates.

Book: Bouton, M. E. (2009). Learning theory. In B. J. Sadock, V. A. Sadock, & P. Ruiz (Eds.), Kaplan & Sadock’s comprehensive textbook of psychiatry (9th ed., Vol. 1, pp. 647–658). New York, NY: Lippincott Williams & Wilkins.

Book: Domjan, M. (2010). The principles of learning and behavior (6th ed.). Belmont, CA: Wadsworth.

Video: Albert Bandura discusses the Bobo Doll Experiment.
https://www.youtube.com/watch?v=eqNaLerMNOE

Discussion Questions

1. Describe three examples of Pavlovian (classical) conditioning that you have seen in your own behavior, or that of your friends or family, in the past few days.
2. Describe three examples of instrumental (operant) conditioning that you have seen in your own behavior, or that of your friends or family, in the past few days.
3. Drugs can be potent reinforcers. Discuss how Pavlovian conditioning and instrumental conditioning can work together to influence drug taking.
4. In the modern world, processed foods are highly available and have been engineered to be highly palatable and reinforcing. Discuss how Pavlovian and instrumental conditioning can work together to explain why people often eat too much.
5. How does blocking challenge the idea that pairings of a CS and US are sufficient to cause Pavlovian conditioning? What is important in creating Pavlovian learning?
6. How does the reinforcer devaluation effect challenge the idea that reinforcers merely “stamp in” the operant response? What does the effect tell us that animals actually learn in operant conditioning?
7. With regards to social learning do you think people learn violence from observing violence

in movies? Why or why not?

8. What do you think you have learned through social learning? Who are your social models?

Vocabulary

Blocking
In classical conditioning, the finding that no conditioning occurs to a stimulus if it is combined with a previously conditioned stimulus during conditioning trials. Suggests that information, surprise value, or prediction error is important in conditioning.

Categorize
To sort or arrange different items into classes or categories.

Classical conditioning
The procedure in which an initially neutral stimulus (the conditioned stimulus, or CS) is paired with an unconditioned stimulus (or US). The result is that the conditioned stimulus begins to elicit a conditioned response (CR). Classical conditioning is nowadays considered important as both a behavioral phenomenon and as a method to study simple associative learning. Same as Pavlovian conditioning.

Conditioned compensatory response
In classical conditioning, a conditioned response that opposes, rather than is the same as, the unconditioned response. It functions to reduce the strength of the unconditioned response. Often seen in conditioning when drugs are used as unconditioned stimuli.

Conditioned response (CR)
The response that is elicited by the conditioned stimulus after classical conditioning has taken place.

Conditioned stimulus (CS)
An initially neutral stimulus (like a bell, light, or tone) that elicits a conditioned response after it has been associated with an unconditioned stimulus.

Context
Stimuli that are in the background whenever learning occurs. For instance, the Skinner box or room in which learning takes place is the classic example of a context. However, “context” can also be provided by internal stimuli, such as the sensory effects of drugs (e.g., being under the influence of alcohol has stimulus properties that provide a context) and mood states (e.g., being happy or sad). It can also be provided by a specific period in time—the passage of time is sometimes said to change the “temporal context.”

Discriminative stimulus
In operant conditioning, a stimulus that signals whether the response will be reinforced. It is said to "set the occasion" for the operant response.

Extinction
Decrease in the strength of a learned behavior that occurs when the conditioned stimulus is presented without the unconditioned stimulus (in classical conditioning) or when the behavior is no longer reinforced (in instrumental conditioning). The term describes both the procedure (the US or reinforcer is no longer presented) as well as the result of the procedure (the learned response declines). Behaviors that have been reduced in strength through extinction are said to be "extinguished."

Fear conditioning
A type of classical or Pavlovian conditioning in which the conditioned stimulus (CS) is associated with an aversive unconditioned stimulus (US), such as a foot shock. As a consequence of learning, the CS comes to evoke fear. The phenomenon is thought to be involved in the development of anxiety disorders in humans.

Goal-directed behavior
Instrumental behavior that is influenced by the animal's knowledge of the association between the behavior and its consequence and the current value of the consequence. Sensitive to the reinforcer devaluation effect.

Habit
Instrumental behavior that occurs automatically in the presence of a stimulus and is no longer influenced by the animal's knowledge of the value of the reinforcer. Insensitive to the reinforcer devaluation effect.

Instrumental conditioning
Process in which animals learn about the relationship between their behaviors and their consequences. Also known as operant conditioning.

Law of effect
The idea that instrumental or operant responses are influenced by their effects. Responses that are followed by a pleasant state of affairs will be strengthened and those that are followed by discomfort will be weakened. Nowadays, the term refers to the idea that operant or instrumental behaviors are lawfully controlled by their consequences.

Observational learning

Learning by observing the behavior of others.

Operant
A behavior that is controlled by its consequences. The simplest example is the rat's lever-pressing, which is controlled by the presentation of the reinforcer.

Operant conditioning
See instrumental conditioning.

Pavlovian conditioning
See classical conditioning.

Prediction error
When the outcome of a conditioning trial is different from that which is predicted by the conditioned stimuli that are present on the trial (i.e., when the US is surprising). Prediction error is necessary to create Pavlovian conditioning (and associative learning generally). As learning occurs over repeated conditioning trials, the conditioned stimulus increasingly predicts the unconditioned stimulus, and prediction error declines. Conditioning works to correct or reduce prediction error.

Preparedness
The idea that an organism's evolutionary history can make it easy to learn a particular association. Because of preparedness, you are more likely to associate the taste of tequila, and not the circumstances surrounding drinking it, with getting sick. Similarly, humans are more likely to associate images of spiders and snakes than flowers and mushrooms with aversive outcomes like shocks.

Punisher
A stimulus that decreases the strength of an operant behavior when it is made a consequence of the behavior.

Quantitative law of effect
A mathematical rule that states that the effectiveness of a reinforcer at strengthening an operant response depends on the amount of reinforcement earned for all alternative behaviors. A reinforcer is less effective if there is a lot of reinforcement in the environment for other behaviors.

Reinforcer
Any consequence of a behavior that strengthens the behavior or increases the likelihood that

it will be performed it again.

Reinforcer devaluation effect
The finding that an animal will stop performing an instrumental response that once led to a reinforcer if the reinforcer is separately made aversive or undesirable.

Renewal effect
Recovery of an extinguished response that occurs when the context is changed after extinction. Especially strong when the change of context involves return to the context in which conditioning originally occurred. Can occur after extinction in either classical or instrumental conditioning.

Social Learning Theory
The theory that people can learn new responses and behaviors by observing the behavior of others.

Social models
Authorities that are the targets for observation and who model behaviors.

Spontaneous recovery
Recovery of an extinguished response that occurs with the passage of time after extinction. Can occur after extinction in either classical or instrumental conditioning.

Stimulus control
When an operant behavior is controlled by a stimulus that precedes it.

Taste aversion learning
The phenomenon in which a taste is paired with sickness, and this causes the organism to reject—and dislike—that taste in the future.

Unconditioned response (UR)
In classical conditioning, an innate response that is elicited by a stimulus before (or in the absence of) conditioning.

Unconditioned stimulus (US)
In classical conditioning, the stimulus that elicits the response before conditioning occurs.

Vicarious reinforcement
Learning that occurs by observing the reinforcement or punishment of another person.

References

Balleine, B. W. (2005). Neural basis of food-seeking: Affect, arousal, and reward in corticostratolimbic circuits. *Physiology & Behavior, 86*, 717–730.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall

Bandura, A., Ross, D., Ross, S (1963). Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology 66*(1), 3 - 11.

Bandura, A.; Ross, D.; Ross, S. A. (1961). "Transmission of aggression through the imitation of aggressive models". *Journal of Abnormal and Social Psychology 63*(3), 575–582.

Bernstein, I. L. (1991). Aversion conditioning in response to cancer and cancer treatment. *Clinical Psychology Review, 11*, 185–191.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory, 11*, 485–494.

Colwill, R. M., & Rescorla, R. A. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *The psychology of learning and motivation*, (Vol. 20, pp. 55–104). New York, NY: Academic Press.

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*, 5–27.

Dickinson, A., & Balleine, B. W. (1994). Motivational control of goal-directed behavior. *Animal Learning & Behavior, 22*, 1–18.

Fanselow, M. S., & Poulos, A. M. (2005). The neuroscience of mammalian associative learning. *Annual Review of Psychology, 56*, 207–234.

Herrnstein, R. J. (1970). On the law of effect. Journal of the Experimental Analysis of Behavior, 13, 243–266.

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes, 30*, 104–117.

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York, NY: Appleton-Century-Crofts.

Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It’s not what you thought it was. *American Psychologist, 61*, 10–26.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52*, 111–139.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Scalera, G., & Bavieri, M. (2009). Role of conditioned taste aversion on the side effects of chemotherapy in cancer patients. In S. Reilly & T. R. Schachtman (Eds.), *Conditioned taste aversion: Behavioral and neural processes* (pp. 513–541). New York, NY: Oxford University Press.

Siegel, S. (1989). Pharmacological conditioning and drug effects. In A. J. Goudie & M. Emmett-Oglesby (Eds.), *Psychoactive drugs* (pp. 115–180). Clifton, NY: Humana Press.

Siegel, S., Hinson, R. E., Krank, M. D., & McCully, J. (1982). Heroin "overdose" death: Contribution of drug associated environmental cues. *Science, 216*, 436–437.

Spreat, S., & Spreat, S. R. (1982). Learning principles. In V. Voith & P. L. Borchelt (Eds.), *Veterinary clinics of North America: Small animal practice* (pp. 593–606). Philadelphia, PA: W. B. Saunders.

Thompson, R. F., & Steinmetz, J. E. (2009). The role of the cerebellum in classical conditioningof discrete behavioral responses. *Neuroscience, 162*, 732–755.

Timberlake, W. L. (2001). Motivational modes in behavior systems. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 155–210). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Wasserman, E. A. (1995). The conceptual abilities of pigeons. *American Scientist, 83*, 246–255.

22
Forgetting and Amnesia

Nicole Dudukovic & Brice Kuhl

This module explores the causes of everyday forgetting and considers pathological forgetting in the context of amnesia. Forgetting is viewed as an adaptive process that allows us to be efficient in terms of the information we retain.

Learning Objectives

- Identify five reasons we forget and give examples of each.
- Describe how forgetting can be viewed as an adaptive process.
- Explain the difference between anterograde and retrograde amnesia.

Introduction

Chances are that you have experienced memory lapses and been frustrated by them. You may have had trouble remembering the definition of a key term on an exam or found yourself unable to recall the name of an actor from one of your favorite TV shows. Maybe you forgot to call your aunt on her birthday or you routinely forget where you put your cell phone. Oftentimes, the bit of information we are searching for comes back to us, but sometimes it does not. Clearly, forgetting seems to be a natural part of life. Why do we forget? And is forgetting always a bad thing?

Causes of Forgetting

Forgetting can often be obnoxious or even embarrassing. But as we explore this module, you'll learn that forgetting is important and necessary for everyday functionality. [Image: jazbeck, https://goo.gl/nkRrJy, CC BY 2.0, https://goo.gl/BRvSA7]

One very common and obvious reason why you cannot remember a piece of information is because you did not learn it in the first place. If you fail to encode information into memory, you are not going to remember it later on. Usually, **encoding** failures occur because we are distracted or are not paying attention to specific details. For example, people have a lot of trouble recognizing an actual penny out of a set of drawings of very similar pennies, or lures, even though most of us have had a lifetime of experience handling pennies (Nickerson & Adams, 1979). However, few of us have studied the features of a penny in great detail, and since we have not attended to those details, we fail to recognize them later. Similarly, it has been well documented that distraction during learning impairs later memory (e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996). Most of the time this is not problematic, but in certain situations, such as when you are studying for an exam, failures to encode due to distraction can have serious repercussions.

Another proposed reason why we forget is that memories fade, or **decay**, over time. It has been known since the pioneering work of Hermann Ebbinghaus (1885/1913) that as time passes, memories get harder to recall. Ebbinghaus created more than 2,000 nonsense syllables, such as *dax*, *bap*, and *rif*, and studied his own memory for them, learning as many as 420 lists of 16 nonsense syllables for one experiment. He found that his memories diminished as time passed, with the most forgetting happening early on after learning. His observations and subsequent research suggested that if we do not rehearse a memory and the neural representation of that memory is not reactivated over a long period of time, the memory representation may disappear entirely or fade to the point where it can no longer be accessed. As you might imagine, it is hard to definitively prove that a memory has decayed as opposed to it being inaccessible for another reason. Critics argued that forgetting must be due to processes other than simply the passage of time, since disuse of a memory does not always guarantee forgetting (McGeoch, 1932). More recently, some memory theorists have proposed that recent memory traces may be degraded or disrupted by new experiences (Wixted, 2004). Memory traces need to be **consolidated**, or transferred from the hippocampus to more durable representations in the cortex, in order for them to last (McGaugh, 2000). When the consolidation process is interrupted by the encoding of other experiences, the memory trace for the original experience does not get fully developed and thus is forgotten.

At times, we will completely blank on something we're certain we've learned - people we went to school with years ago for example. However, once we get the right retrieval cue (a name perhaps), the memory (faces or experiences) rushes back to us like it was there all along. [Image: sbhsclass84, https://goo.gl/sHZyQI, CC BY-SA 2.0, https://goo.gl/rxiUsF]

Both encoding failures and decay account for more permanent forms of forgetting, in which the memory trace does not exist, but forgetting may also occur when a memory exists yet we temporarily cannot access it. This type of forgetting may occur when we lack the appropriate **retrieval** cues for bringing the memory to mind. You have probably had the frustrating experience of forgetting your password for an online site. Usually, the password has not been permanently forgotten; instead, you just need the right reminder to remember what it is. For example, if your password was "pizza0525," and you received the password hints "favorite food" and "Mom's birthday," you would easily be able to retrieve it. Retrieval hints can bring back to mind seemingly forgotten memories (Tulving & Pearlstone, 1966). One real-life illustration of the importance of retrieval cues comes from a study showing that whereas people have difficulty recalling the names of high school classmates years after graduation, they are easily able to recognize the names and match them to the appropriate faces (Bahrick, Bahrick, & Wittinger, 1975). The names are powerful enough retrieval cues that they bring back the memories of the faces that went with them. The fact that the presence of the right retrieval cues is critical for remembering adds to the difficulty in proving that a memory is permanently forgotten as opposed to temporarily unavailable.

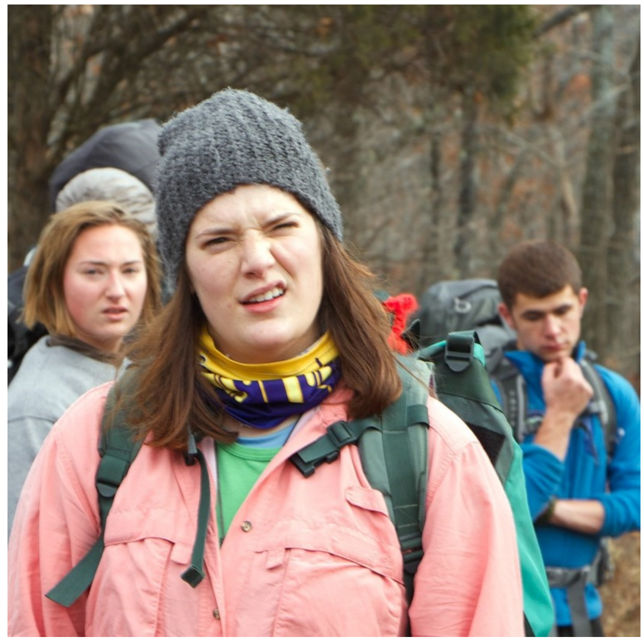
Retrieval failures can also occur because other memories are blocking or getting in the way of recalling the desired memory. This blocking is referred to as **interference**. For example, you may fail to remember the name of a town you visited with your family on summer vacation because the names of other towns you visited on that trip or on other trips come to mind instead. Those memories then prevent the desired memory from being retrieved. Interference is also relevant to the example of forgetting a password: passwords that we have used for other websites may come to mind and interfere with our ability to retrieve the desired password. Interference can be either proactive, in which old memories block the learning of new related memories, or retroactive, in which new memories block the retrieval of old related memories. For both types of interference, competition between memories seems to be key (Mensink & Raaijmakers, 1988). Your memory for a town you visited on vacation is unlikely to interfere with your ability to remember an Internet password, but it is likely to interfere with your ability to remember a different town's name. Competition between memories can also lead to forgetting in a different way. Recalling a desired memory in the face of competition may result in the inhibition of related, competing memories (Levy & Anderson, 2002). You may have difficulty recalling the name of Kennebunkport, Maine, because other Maine towns, such as Bar Harbor, Winterport, and Camden, come to mind instead. However, if you are able to recall Kennebunkport despite strong competition from the other towns, this may actually change the competitive landscape, weakening memory for those other towns' names, leading to forgetting of them instead.

Box 1. Five Impediments to Remembering

1. Encoding failures - we don't learn the information in the first place
2. Decay - memories fade over time
3. Inadequate retrieval cues - we lack sufficient reminders
4. Interference - other memories get in the way
5. Trying not to remember - we deliberately attempt to keep things out of mind

Finally, some memories may be forgotten because *we deliberately attempt to keep them out of mind*. Over time, by actively trying not to remember an event, we can sometimes successfully keep the undesirable memory from being retrieved either by inhibiting the undesirable memory or generating diversionary thoughts (Anderson & Green, 2001). Imagine that you slipped and fell in your high school cafeteria during lunch time, and everyone at the surrounding tables laughed at you. You would likely wish to avoid thinking about that event and might try to prevent it from coming to mind. One way that you could accomplish this is by thinking of other, more positive, events that are associated with the cafeteria. Eventually, this memory may be suppressed to the point that it would only be retrieved with great difficulty (Hertel & Calcaterra, 2005).

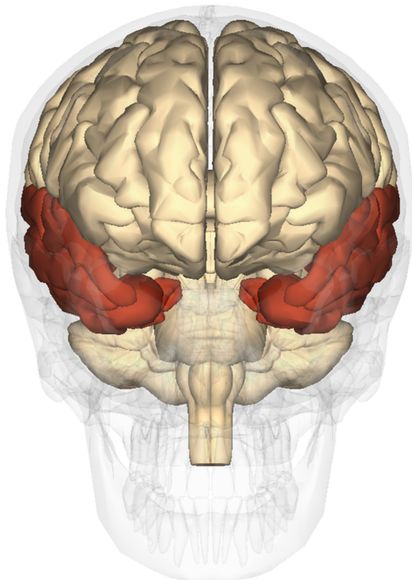
Adaptive Forgetting



Could you imagine being unable to forget every path you have taken while hiking? Each new trip, you would be walking around the forest for days, incapable of distinguishing today's path from the prior ones. [Image: Dan Trew, https://goo.gl/8fJWWE, CC BY-SA 2.0, https://goo.gl/rxiUsF]

We have explored five different causes of forgetting. Together they can account for the day-to-day episodes of forgetting that each of us experience. Typically, we think of these episodes in a negative light and view forgetting as a memory failure. Is forgetting ever good? Most people would reason that forgetting that occurs in response to a deliberate attempt to keep an event out of mind is a good thing. No one wants to be constantly reminded of falling on their face in front of all of their friends. However, beyond that, it can be argued that forgetting is adaptive, allowing us to be efficient and hold onto only the most relevant memories (Bjork, 1989; Anderson & Milson, 1989). Shereshevsky, or "S," the mnemonist studied by Alexander Luria (1968), was a man who almost never forgot. His memory appeared to be virtually limitless. He could memorize a table of 50 numbers in under 3 minutes and recall the numbers in rows, columns, or diagonals with ease. He could recall lists of words and passages that he had memorized over a decade before. Yet Shereshevsky found it difficult to function in his everyday life because he was constantly distracted by a flood of details and associations that sprung to mind. His case history suggests that remembering everything is not always a good thing. You may occasionally have trouble remembering where you parked your car, but imagine if every time you had to find your car, every single former parking space came to mind. The task would become impossibly difficult to sort through all of those irrelevant memories. Thus, forgetting is adaptive in that it makes us more efficient. The price of that efficiency is those moments when our memories seem to fail us (Schacter, 1999).

Amnesia



Patients with damage to the temporal lobes may experience anterograde amnesia and/or retrograde amnesia. [Image: en: Anatomography, https://goo.gl/ALPAu6, CC BY-SA 2.1 JP, https://goo.gl/BDF2Z4]

Clearly, remembering everything would be maladaptive, but what would it be like to remember nothing? We will now consider a profound form of forgetting called amnesia that is distinct from more ordinary forms of forgetting. Most of us have had exposure to the concept of amnesia through popular movies and television. Typically, in these fictionalized portrayals of amnesia, a character suffers some type of blow to the head and suddenly has no idea who they are and can no longer recognize their family or remember any events from their past. After some period of time (or another blow to the head), their memories come flooding back to them. Unfortunately, this portrayal of amnesia is not very accurate. What does amnesia typically look like?

The most widely studied amnesic patient was known by his initials H. M. (Scoville & Milner, 1957). As a teenager, H. M. suffered from severe epilepsy, and in 1953, he underwent surgery to have both of his medial temporal lobes removed to relieve his epileptic seizures. The **medial temporal lobes** encompass the hippocampus and surrounding cortical tissue. Although the surgery was successful in reducing H. M.'s seizures and his general intelligence was preserved, the surgery left H. M. with a profound and permanent memory deficit. From the time of his surgery until his death in 2008, H. M. was unable to learn new information, a memory impairment called **anterograde amnesia**. H. M. could not remember any event that occurred since his surgery, including highly significant ones, such as the death of his father. He could not remember a conversation he had a few minutes prior or recognize the face of someone who had visited him that same day. He could keep information in his short-term, or working, memory, but when his attention turned to something else, that information was lost for good. It is important to note that H. M.'s memory impairment was restricted to **declarative memory**, or conscious memory for facts and events. H. M. could learn new motor skills and showed improvement on motor tasks even in the absence of any memory for having performed the task before (Corkin, 2002).

In addition to anterograde amnesia, H. M. also suffered from **temporally graded retrograde amnesia**. **Retrograde amnesia** refers to an inability to retrieve old memories that occurred before the onset of amnesia. Extensive retrograde amnesia in the absence of anterograde amnesia is very rare (Kopelman, 2000). More commonly, retrograde amnesia co-occurs with anterograde amnesia and shows a temporal gradient, in which memories closest in time to the onset of amnesia are lost, but more remote memories are retained (Hodges, 1994). In the case of H. M., he could remember events from his childhood, but he could not remember events that occurred a few years before the surgery.

Amnesiac patients with damage to the hippocampus and surrounding medial temporal lobes typically manifest a similar clinical profile as H. M. The degree of anterograde amnesia and retrograde amnesia depend on the extent of the medial temporal lobe damage, with greater damage associated with a more extensive impairment (Reed & Squire, 1998). Anterograde amnesia provides evidence for the role of the hippocampus in the formation of long-lasting declarative memories, as damage to the hippocampus results in an inability to create this type of new memory. Similarly, temporally graded retrograde amnesia can be seen as providing further evidence for the importance of memory consolidation (Squire & Alvarez, 1995). A memory depends on the hippocampus until it is consolidated and transferred into a more durable form that is stored in the cortex. According to this theory, an amnesiac patient like H. M. could remember events from his remote past because those memories were fully consolidated and no longer depended on the hippocampus.

The classic amnesiac syndrome we have considered here is sometimes referred to as organic amnesia, and it is distinct from functional, or dissociative, amnesia. Functional amnesia involves a loss of memory that cannot be attributed to brain injury or any obvious brain disease and is typically classified as a mental disorder rather than a neurological disorder (Kihlstrom, 2005). The clinical profile of dissociative amnesia is very different from that of patients who suffer from amnesia due to brain damage or deterioration. Individuals who experience **dissociative amnesia** often have a history of trauma. Their amnesia is retrograde, encompassing autobiographical memories from a portion of their past. In an extreme version of this disorder, people enter a dissociative fugue state, in which they lose most or all of their autobiographical memories and their sense of personal identity. They may be found wandering in a new location, unaware of who they are and how they got there. Dissociative amnesia is controversial, as both the causes and existence of it have been called into question. The memory loss associated with dissociative amnesia is much less likely to be permanent than it is in organic amnesia.

Conclusion

Just as the case study of the mnemonist Shereshevsky illustrates what a life with a near perfect memory would be like, amnesiac patients show us what a life without memory would be like. Each of the mechanisms we discussed that explain everyday forgetting—encoding failures, decay, insufficient retrieval cues, interference, and intentional attempts to forget—help to keep us highly efficient, retaining the important information and for the most part, forgetting the unimportant. Amnesiac patients allow us a glimpse into what life would be like if we suffered from profound forgetting and perhaps show us that our everyday lapses in memory are not so bad after all.

Outside Resources

Web: Brain Case Study: Patient HM
https://bigpictureeducation.com/brain-case-study-patient-hm

Web: Self-experiment, Penny demo
http://www.indiana.edu/~p1013447/dictionary/penny.htm

Web: The Man Who Couldn’t Remember
http://www.pbs.org/wgbh/nova/body/corkin-hm-memory.html

Discussion Questions

1. Is forgetting good or bad? Do you agree with the authors that forgetting is an adaptive process? Why or why not?
2. Can we ever prove that something is forgotten? Why or why not?
3. Which of the five reasons for forgetting do you think explains the majority of incidences of forgetting? Why?
4. How is real-life amnesia different than amnesia that is portrayed on TV and in film?

Vocabulary

Anterograde amnesia
Inability to form new memories for facts and events after the onset of amnesia.

Consolidation
Process by which a memory trace is stabilized and transformed into a more durable form.

Decay
The fading of memories with the passage of time.

Declarative memory
Conscious memories for facts and events.

Dissociative amnesia
Loss of autobiographical memories from a period in the past in the absence of brain injury or disease.

Encoding
Process by which information gets into memory.

Interference
Other memories get in the way of retrieving a desired memory

Medial temporal lobes
Inner region of the temporal lobes that includes the hippocampus.

Retrieval
Process by which information is accessed from memory and utilized.

Retrograde amnesia
Inability to retrieve memories for facts and events acquired before the onset of amnesia.

Temporally graded retrograde amnesia
Inability to retrieve memories from just prior to the onset of amnesia with intact memory for more remote events.

References

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703–719.

Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, 410, 366–369.

Bahrick, H. P., Bahrick, P. O., & Wittinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, 104, 54–75.

Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of Memory and Consciousness* (pp. 309– 330). Hillsdale, NJ: Erlbaum.

Corkin, S. (2002). What's new with the amnesic patient H. M.? *Nature Reviews Neuroscience*, 3, 153–160.

Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125, 159–180.

Ebbinghaus, H. (1913). *Memory. A contribution to experimental psychology*. New York: Teachers College/Columbia University (Engl. ed.). (Original work published in 1885.)

Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, 12, 484–489.

Hodges, J. R. (1994). Retrograde amnesia. In A. Baddeley, B. A. Wilson, & F. Watts (Eds.), *Handbook of Memory Disorders* (pp. 81–107). New York: Wiley.

Kihlstrom, J. F. (2005). Dissociative disorders. Annual *Review of Clinical Psychology*, 1, 227– 253.

Kopelman, M. (2000). Focal retrograde amnesia and the attribution of causality: An exceptionally critical review. *Cognitive Neuropsychology*, 17, 585–621.

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, 6, 299–305.

Luria, A. R. (1968). *The mind of a mnemonist: A little book about a vast memory* (L. Solataroff, Trans.). New York: Basic Books.

McGaugh, J. L. (2000). Memory: A century of consolidation. *Science*, 287, 248–251.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Reviews*, 39, 352– 370.

Mensink, G., & Raaijmakers, J. G. (1988). A model for interference and forgetting. *Psychological Review*, 95, 434–455.

Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, 11, 287–307.

Reed, J. M. & Squire, L. R. (1998). Retrograde amnesia for facts and events: Findings from four new cases. *Journal of Neuroscience*, 18, 3943–3954.

Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54, 182–203.

Scoville, W. B. & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology*, Neurosurgery, & Psychiatry, 20, 11–21.

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinions in Neurobiology*, 5, 169–177.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Reviews of Psychology*, 55, 235–269.

Cognition and Language

23
Attention

Frances Friedrich

We use the term “attention“ all the time, but what processes or abilities does that concept really refer to? This module will focus on how attention allows us to select certain parts of our environment and ignore other parts, and what happens to the ignored information. A key concept is the idea that we are limited in how much we can do at any one time. So we will also consider what happens when someone tries to do several things at once, such as driving while using electronic devices.

Learning Objectives

- Understand why selective attention is important and how it can be studied.
- Learn about different models of when and how selection can occur.
- Understand how divided attention or multitasking is studied, and implications of multitasking in situations such as distracted driving.

What is Attention?

Before we begin exploring attention in its various forms, take a moment to consider how you think about the concept. How would you define attention, or how do you use the term? We certainly use the word very frequently in our everyday language: “ATTENTION! USE ONLY AS DIRECTED!” warns the label on the medicine bottle, meaning be alert to possible danger. “Pay attention!” pleads the weary seventh-grade teacher, not warning about danger (with possible exceptions, depending on the teacher) but urging the students to focus on the task at hand.

We may refer to a child who is easily distracted as having an attention disorder, although we also are told that Americans have an attention span of about 8 seconds, down from 12 seconds in 2000, suggesting that we *all* have trouble sustaining concentration for any amount of time (from **www.Statisticbrain.com).** How that number was determined is not clear from the Web site, nor is it clear how attention span in the goldfish—9 seconds!—was measured, but the fact that our average span reportedly is less than that of a goldfish is intriguing, to say the least.

William James wrote extensively about attention in the late 1800s. An often quoted passage (James, 1890/1983) beautifully captures how intuitively obvious the concept of attention is, while it remains very difficult to define in measurable, concrete terms:

> Everyone 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. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others. (pp. 381–382)

Are you reading these words right here right now? If so, it's only because you directed your attention toward them. [Image: CC BY 2.0, https://goo.gl/BRvSA7]

Notice that this description touches on the conscious nature of attention, as well as the notion that what is in consciousness is often controlled voluntarily but can also be determined by events that capture our attention. Implied in this description is the idea that we seem to have a **limited capacity** for information processing, and that we can only attend to or be consciously aware of a small amount of information at any given time.

Many aspects of attention have been studied in the field of psychology. In some respects, we define different types of attention by the nature of the task used to study it. For example, a crucial issue in World War II was how long an individual could remain highly alert and accurate while watching a radar screen for enemy planes, and this problem led psychologists to study how attention works under such conditions. When watching for a rare event, it is easy to allow concentration to lag. (This a continues to be a challenge today for TSA agents, charged with looking at images of the contents of your carry-on items in search of knives, guns, or shampoo bottles larger than 3 oz.) Attention in the context of this type of search task refers to the level of *sustained attention* or *vigilance* one can maintain. In contrast, **divided attention** tasks allow us to determine how well individuals can attend to many sources of information at once. *Spatial attention* refers specifically to how we focus on one part of our environment and how we move attention to other locations in the environment. These are all examples of different aspects of attention, but an implied element of most of these ideas is the concept of **selective attention**; some information is attended to while other information is intentionally blocked out. This module will focus on important issues in selective and divided attention, addressing these questions:

- Can we pay attention to several sources of information at once, or do we have a limited capacity for information?
- How do we select what to pay attention to?
- What happens to information that we try to ignore?
- Can we learn to divide attention between multiple tasks?

Selective Attention

The Cocktail Party

Selective attention is *the ability to select certain stimuli in the environment to process, while ignoring distracting information*. One way to get an intuitive sense of how attention works is to consider situations in which attention is used. A party provides an excellent example for our purposes. Many people may be milling around, there is a dazzling variety of colors and sounds and smells, the buzz of many conversations is striking. There are so many conversations going on; how is it possible to select just one and follow it? You don't have to be looking at the person talking; you may be listening with great interest to some gossip while pretending not to hear. However, once you are engaged in conversation with someone, you quickly become aware that

Beyond just hearing your name from the clamor at a party, other words or concepts, particularly unusual or significant ones to you, can also snag your attention. [Image: Ross, https://goo.gl/TVDfTn, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF] you cannot also listen to other conversations at the same time. You also are probably *not* aware of how tight your shoes feel or of the smell of a nearby flower arrangement. On the other hand, if someone behind you mentions your name, you typically notice it immediately and may start attending to that (much more interesting) conversation. This situation highlights an interesting set of observations. We have an amazing ability to select and track one voice, visual object, etc., even when a million things are competing for our attention, but at the same time, we seem to be limited in how much we can attend to at one time, which in turn suggests that attention is crucial in selecting what is important. How does it all work?

Dichotic Listening Studies

This cocktail party scenario is the quintessential example of selective attention, and it is essentially what some early researchers tried to replicate under controlled laboratory conditions as a starting point for understanding the role of attention in perception (e.g., Cherry, 1953; Moray, 1959). In particular, they used **dichotic listening** and **shadowing** tasks to evaluate the selection process. Dichotic listening simply refers to the situation when two messages are presented simultaneously to an individual, with one message in each ear. In order to control which message the person attends to, the individual is asked to repeat back or “shadow” one of the messages as he hears it. For example, let’s say that a story about a camping trip is presented to John’s left ear, and a story about Abe Lincoln is presented to his right ear. The typical dichotic listening task would have John repeat the story presented to one ear as he hears it. Can he do that without being distracted by the information in the other ear?

People can become pretty good at the shadowing task, and they can easily report the content of the message that they attend to. But what happens to the ignored message? Typically, people can tell you if the ignored message was a man’s or a woman’s voice, or other physical characteristics of the speech, but they cannot tell you what the message was about. In fact, many studies have shown that people in a shadowing task were not aware of a change in the language of the message (e.g., from English to German; Cherry, 1953), and they didn't even notice when the same word was repeated in the unattended ear more than 35 times (Moray, 1959)! Only the basic physical characteristics, such as the pitch of the unattended message, could be reported.

On the basis of these types of experiments, it seems that we can answer the first question about how much information we can attend to very easily: not very much. We clearly have a limited capacity for processing information for meaning, making the selection process all the more important. The question becomes: How does this selection process work?

Models of Selective Attention

Broadbent's Filter Model. Many researchers have investigated how selection occurs and what happens to ignored information. Donald Broadbent was one of the first to try to characterize the selection process. His Filter Model was based on the dichotic listening tasks described above as well as other types of experiments (Broadbent, 1958). He found that people select information on the basis of *physical features*: the sensory channel (or ear) that a message was coming in, the pitch of the voice, the color or font of a visual message. People seemed vaguely aware of the physical features of the unattended information, but had no knowledge of the meaning. As a result, Broadbent argued that selection occurs *very early*, with no additional processing for the unselected information. A flowchart of the model might look like this:

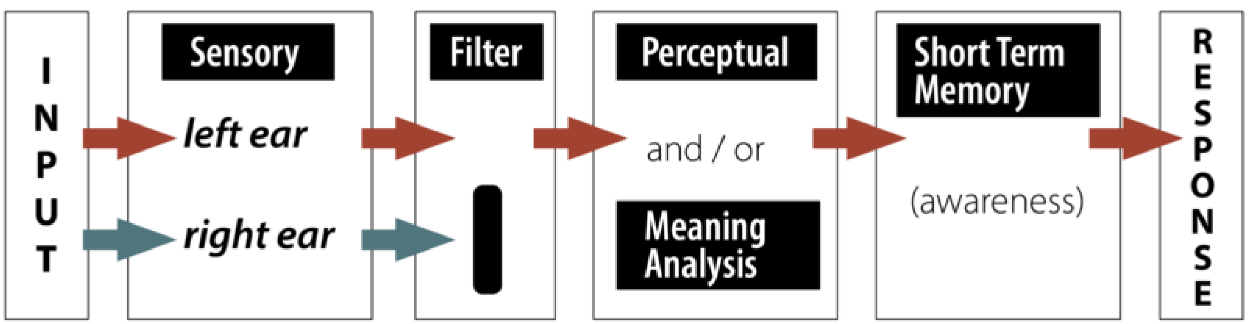


Figure 1: This figure shows information going in both the left and right ears. Some basic sensory information, such as pitch, is processed, but the filter only allows the information from one ear to be processed further. Only the information from the left ear is transferred to short-term memory (STM) and conscious awareness, and then further processed for meaning. That means that the ignored information never makes it beyond a basic physical analysis.

Treisman' s Attenuation Model

Broadbent's model makes sense, but if you think about it you already know that it cannot account for all aspects of the Cocktail Party Effect. What doesn't fit? The fact is that you tend to hear your own name when it is spoken by someone, even if you are deeply engaged in a conversation. We mentioned earlier that people in a shadowing experiment were unaware of a word in the unattended ear that was repeated many times—and yet many people noticed their own name in the unattended ear even it occurred only once.

Anne Treisman (1960) carried out a number of dichotic listening experiments in which she presented two different stories to the two ears. As usual, she asked people to shadowthe message in one ear. As the stories progressed, however, she switched the stories to the opposite ears. Treisman found that individuals spontaneously followed the story, or the content of the message, when it shifted from the left ear to the right ear. Then they realized they were shadowing the wrong ear and switched back.

Results like this, and the fact that you tend to hear meaningful information even when you aren't paying attention to it, suggest that we *do* monitor the unattended information to some degree on the basis of its meaning. Therefore, the filter theory can't be right to suggest that unattended information is completely blocked at the sensory analysis level. Instead, Treisman suggested that selection *starts* at the physical or perceptual level, but that the unattended information is not blocked completely, it is just weakened or *attenuated*. As a result, highly meaningful or pertinent information in the unattended ear will get through the filter for further processing at the level of meaning. The figure below shows information going in both ears, and in this case there is no filter that completely blocks nonselected information. Instead, selection of the left ear information strengthens that material, while the nonselected information in the right ear is weakened. However, if the preliminary analysis shows that the nonselected information is especially pertinent or meaningful (such as your own name), then the Attenuation Control will instead strengthen the more meaningful information.

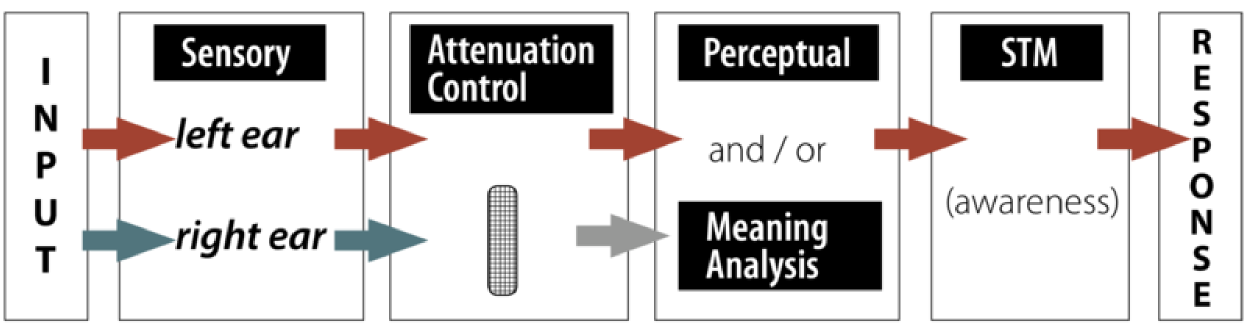


Figure 2

Late Selection Models

Other selective attention models have been proposed as well. A *late selection* or *response selection* model proposed by Deutsch and Deutsch (1963) suggests that all information in the unattended ear is processed on the basis of meaning, not just the selected or highly pertinent information. However, only the information that is relevant for the task response gets into conscious awareness. This model is consistent with ideas of subliminal perception; in other words, that you don't have to be aware of or attending a message for it to be fully processed for meaning.

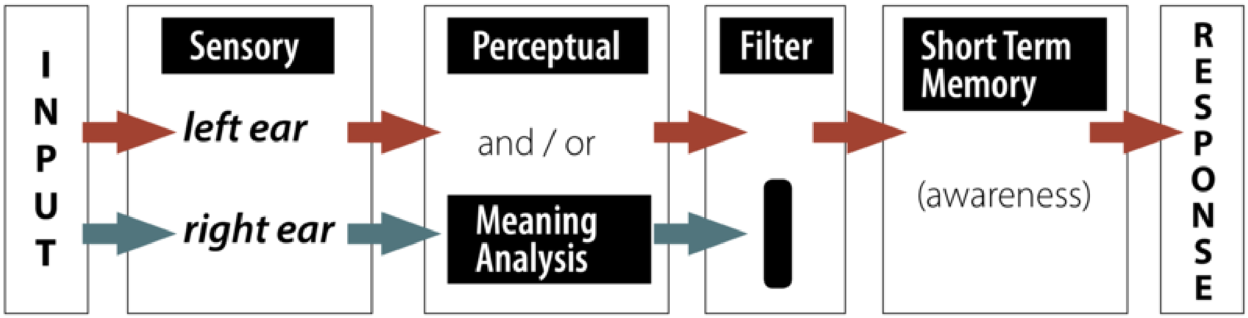


Figure 3

You might notice that this figure looks a lot like that of the Early Selection model—only the location of the selective filter has changed, with the assumption that analysis of meaning occurs *before* selection occurs, but only the selected information becomes conscious.

Multimode Model

Why did researchers keep coming up with different models? Because no model really seemed to account for all the data, some of which indicates that nonselected information is blocked completely, whereas other studies suggest that it can be processed for meaning. The multimode model addresses this apparent inconsistency, suggesting that the stage at which selection occurs can change depending on the task. Johnston and Heinz (1978) demonstrated that under some conditions, we can select what to attend to at a very early stage and we do not process the content of the unattended message very much at all. Analyzing physical information, such as attending to information based on whether it is a male or female voice, is relatively easy; it occurs automatically, rapidly, and doesn't take much effort. Under the right conditions, we can select what to attend to on the basis of the meaning of the messages. However, the late selection option—processing the content of all messages before selection—is more difficult and requires more effort. The benefit, though, is that we have the flexibility to change how we deploy our attention depending upon what we are trying to accomplish, which is one of the greatest strengths of our cognitive system.

This discussion of selective attention has focused on experiments using auditory material, but the same principles hold for other perceptual systems as well. Neisser (1979) investigated some of the same questions with visual materials by superimposing two semi-transparent video clips and asking viewers to attend to just one series of actions. As with the auditory materials, viewers often were unaware of what went on in the other clearly visible video. Twenty years later, Simons and Chabris (1999) explored and expanded these findings using similar techniques, and triggered a flood of new work in an area referred to as inattentional blindness. We touch on those ideas below, and you can also refer to another Noba Module, **Failures of Awareness: The Case of Inattentional Blindness** for a more complete discussion.

Focus Topic 1: Subliminal Perception

The idea of subliminal perception—that stimuli presented below the threshold for awareness can influence thoughts, feelings, or actions—is a fascinating and kind of creepy one. Can messages you are unaware of, embedded in movies or ads or the music playing in the grocery store, really influence what you buy? Many such claims of the power of subliminal perception have been made. One of the most famous came from a market researcher who claimed that the message "Eat Popcorn" briefly flashed throughout a movie increased popcorn sales by more than 50%, although he later admitted that the study was made up (Merikle, 2000). Psychologists have worked hard to investigate whether this is a valid phenomenon. Studying subliminal perception is more difficult than it might seem, because of the difficulty of establishing what the threshold for consciousness is or of even determining what type of threshold is important; for example, Cheesman and Merikle (1984, 1986) make an important distinction between objective and subjective thresholds. The bottom line is that there is some evidence that individuals can be influenced by stimuli they are not aware of, but how complex the stimuli can be or the extent to which unconscious material can affect behavior is not settled (e.g., Bargh & Morsella, 2008; Greenwald, 1992; Merikle, 2000).

Divided Attention and Multitasking

In spite of the evidence of our limited capacity, we all like to think that we can do several things at once. Some people claim to be able to multitask without any problem: reading a textbook while watching television and talking with friends; talking on the phone while playing computer games; texting while driving. The fact is that we sometimes can *seem* to juggle several things at once, but the question remains whether dividing attention in this way impairs performance.

Is it possible to overcome the limited capacity that we experience when engaging in cognitive tasks? We know that with extensive practice, we can acquire skills that do not appear to require conscious attention. As we walk down the street, we don't need to think consciously about what muscle to contract in order to take the next step. Indeed, paying attention to automated skills can lead to a breakdown in performance, or "choking" (e.g., Beilock & Carr, 2001). But what about higher level, more mentally demanding tasks: Is it possible to learn to perform two complex tasks at the same time?

Divided Attention Tasks

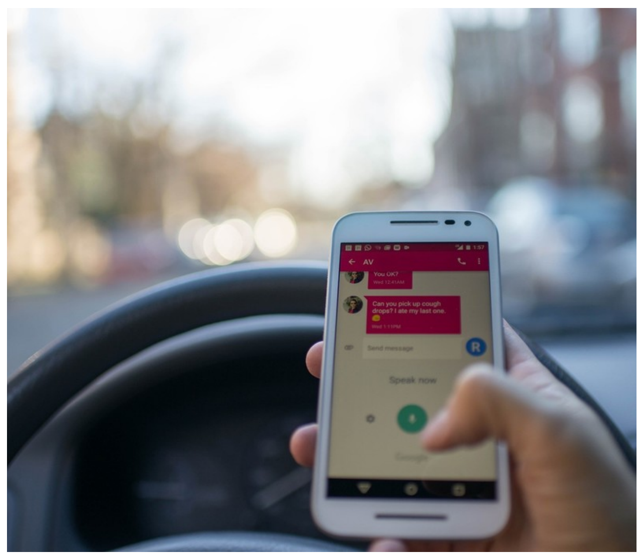
Unless a task is fully automated, some researchers suggest that "multi-tasking" doesn't really exist; you are just rapidly switching your attention back and forth between tasks. [Image: CC0 Public Domain, https://goo.gl/m25gce]

In a classic study that examined this type of divided attention task, two participants were trained to take dictation for spoken words while reading unrelated material for comprehension (Spelke, Hirst, & Neisser, 1976). In divided attention tasks such as these, each task is evaluated separately, in order to determine baseline performance when the individual can allocate as many cognitive resources as necessary to one task at a time. Then performance is evaluated when the two tasks are performed simultaneously. A decrease in performance for either task would suggest that even if attention can be divided or switched between the tasks, the cognitive demands are too great to avoid disruption of performance. (We should note here that divided attention tasks are designed, in principle, to see if two tasks can be carried out simultaneously. A related research area looks at *task switching* and how well we can switch back and forth among different tasks [e.g., Monsell, 2003]. It turns out that switching itself is cognitively demanding and can impair performance.)

The focus of the Spelke et al. (1976) study was whether individuals could learn to perform two relatively complex tasks concurrently, without impairing performance. The participants received plenty of practice—the study lasted 17 weeks and they had a 1-hour session each day, 5 days a week. These participants were able to learn to take dictation for lists of words and read for comprehension without affecting performance in either task, and the authors suggested that perhaps there are not fixed limits on our attentional capacity. However, changing the tasks somewhat, such as reading aloud rather than silently, impaired performance initially, so this multitasking ability may be specific to these well-learned tasks. Indeed, not everyone could learn to perform two complex tasks without performance costs (Hirst, Neisser, & Spelke, 1978), although the fact that some can is impressive.

Distracted Driving

More relevant to our current lifestyles are questions about multitasking while texting or having cell phone conversations. Research designed to investigate, under controlled conditions, multitasking while driving has revealed some surprising results. Certainly there are many possible types of distractions that could impair driving performance, such as applying makeup using the rearview mirror, attempting (usually in vain) to stop the kids in the backseat from fighting, fiddling with the CD player, trying to negotiate a handheld cell phone, a cigarette, and a soda all at once, eating a bowl of cereal while driving (!). But we tend to have a strong sense that we CAN multitask while driving, and cars are being built with more and more technological capabilities that encourage multitasking. How good are we at dividing attention in these cases?



If you look at your phone for just 5 seconds while driving at 55mph, that means you have driven the length of a football field without looking at the road. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Most people acknowledge the distraction caused by texting while driving and the reason seems obvious: Your eyes are off the road and your hands and at least one hand (often both) are engaged while texting. However, the problem is not simply one of occupied hands or eyes, but rather that the *cognitive* demands on our limited capacity systems can seriously impair driving performance (Strayer, Watson, & Drews, 2011). The effect of a cell phone conversation on performance (such as not noticing someone's brake lights or responding more slowly to them) is just as significant when the individual is having a conversation with a hands-free device as with a handheld phone; the same impairments do not occur when listening to the radio or a book on tape (Strayer & Johnston, 2001). Moreover, studies using eye-tracking devices have shown that drivers are less likely to later recognize objects that they *did* look at when using a cell phone while driving (Strayer & Drews, 2007). These findings demonstrate that cognitive distractions such as cell phone conversations can produce inattentional blindness, or a lack of awareness of what is right before your eyes (see also, Simons & Chabris, 1999). Sadly, although we all like to think that we can multitask while driving, in fact the percentage of people who can truly perform cognitive tasks without impairing their driving performance is estimated to be about 2% (Watson & Strayer, 2010).

Summary

It may be useful to think of attention as a mental resource, one that is needed to focus on and fully process important information, especially when there is a lot of distracting “noise” threatening to obscure the message. Our selective attention system allows us to find or track an object or conversation in the midst of distractions. Whether the selection process occurs early or late in the analysis of those events has been the focus of considerable research, and in fact how selection occurs may very well depend on the specific conditions. With respect to divided attention, in general we can only perform one cognitively demanding task at a time, and we may not even be aware of unattended events even though they might seem too obvious to miss (check out some examples in the Outside Resources below). This type of inattention blindness can occur even in well-learned tasks, such as driving while talking on a cell phone. Understanding how attention works is clearly important, even for our everyday lives.

Outside Resources

Video: Here's a wild example of how much we fail to notice when our attention is captured by one element of a scene.
http://www.youtube.com/watch?v=ubNF9QNEQLA&feature=related

Video: Try this test to see how well you can focus on a task in the face of a lot of distraction.
http://www.youtube.com/watch?v=Ahg6qcgoay4&NR=1

Discussion Questions

1. Discuss the implications of the different models of selective attention for everyday life. For instance, what advantages and disadvantages would be associated with being able to filter out all unwanted information at a very early stage in processing? What are the implications of processing all ignored information fully, even if you aren't consciously aware of that information?
2. Think of examples of when you feel you can successfully multitask and when you can’t. Discuss what aspects of the tasks or the situation seem to influence divided attention performance. How accurate do you think you are in judging your own multitasking ability?
3. What are the public policy implications of current evidence of inattentional blindness as a result of distracted driving? Should this evidence influence traffic safety laws? What additional studies of distracted driving would you propose?

Vocabulary

Dichotic listening
An experimental task in which two messages are presented to different ears.

Divided attention
The ability to flexibly allocate attentional resources between two or more concurrent tasks.

Inattentional blindness
The failure to notice a fully visible object when attention is devoted to something else.

Limited capacity
The notion that humans have limited mental resources that can be used at a given time.

Selective attention
The ability to select certain stimuli in the environment to process, while ignoring distracting information.

Shadowing
A task in which the individual is asked to repeat an auditory message as it is presented.

Subliminal perception
The ability to process information for meaning when the individual is not consciously aware of that information.

References

Bargh, J., & Morsella, E. (2008). The unconscious mind. *Perspectives on Psychological Science, 3* (1), 73–79.

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*, 701–725.

Broadbent, D. A. (1958). *Perception and communication*. London, England: Pergamon Press.

Cheesman, J., & Merikle, P. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology, 40*, 343–367.

Cheesman, J., & Merikle, P. (1984). Priming with and without awareness. *Perception and Psychophysics, 36*, 387–395.

Cherry, E. C. (1953). Experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 25*, 975–979.

Deutsch, J. A., & Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological Review, 70*, 80–90.

Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *American Psychologist, 47*, 766–779.

Hirst, W. C., Neisser, U., & Spelke, E. S. (1978). Divided attention. *Human Nature, 1*, 54–61.

James, W. (1983). *The principles of psychology*. Cambridge, MA: Harvard University Press. (Original work published 1890)

Johnston, W. A., & Heinz, S. P. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology: General, 107*, 420–435.

Merikle, P. (2000). Subliminal perception. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 7, pp. 497–499). New York, NY: Oxford University Press.

Monsell, S. (2003). Task switching. *Trends in Cognitive Science, 7*(3), 134–140.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56–60.

Neisser, U. (1979). The control of information pickup in selective looking. In A. D. Pick (Ed.), *Perception and its development: A tribute to Eleanor J. Gibson* (pp. 201–219). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*, 1059–1074.

Spelke, E. S., Hirst, W. C., & Neisser, U. (1976). Skills of divided attention. *Cognition, 4*, 215–250.

Strayer, D. L., & Drews, F. A. (2007). Cell-phone induced inattention blindness. *Current Directions in Psychological Science, 16*, 128–131.

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science, 12*, 462–466.

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011) Cognitive distraction while multitasking in the automobile. In Brian Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 54, pp. 29–58). Burlington, VT: Academic Press.

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12*, 242–248.

Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review, 17*, 479–485.

24
Intelligence

Robert Biswas-Diener

Intelligence is among the oldest and longest studied topics in all of psychology. The development of assessments to measure this concept is at the core of the development of psychological science itself. This module introduces key historical figures, major theories of intelligence, and common assessment strategies related to intelligence. This module will also discuss controversies related to the study of group differences in intelligence.

Learning Objectives

- List at least two common strategies for measuring intelligence.
- Name at least one "type" of intelligence.
- Define intelligence in simple terms.
- Explain the controversy relating to differences in intelligence between groups.

Introduction

Every year hundreds of grade school students converge on Washington, D.C., for the annual Scripps National Spelling Bee. The "bee" is an elite event in which children as young as 8 square off to spell words like "cymotrichous" and "appoggiatura." Most people who watch the bee think of these kids as being "smart" and you likely agree with this description.

What makes a person intelligent? Is it heredity (two of the 2014 contestants in the bee have siblings who have previously won)(National Spelling Bee, 2014a)? Is it interest (the most

A participant in the Scripps National Spelling Bee. [Image: Scripps National Spelling Bee, https://goo.gl/94Hgbm, CC BY-NC 2.0, https://goo.gl/VnKlK8]

frequently listed favorite subject among spelling bee competitors is math)(NSB, 2014b)? In this module we will cover these and other fascinating aspects of **intelligence**. By the end of the module you should be able to define intelligence and discuss some common strategies for measuring intelligence. In addition, we will tackle the politically thorny issue of whether there are differences in intelligence between groups such as men and women.

Defining and Measuring Intelligence

When you think of "smart people" you likely have an intuitive sense of the qualities that make them intelligent. Maybe you think they have a good memory, or that they can think quickly, or that they simply know a whole lot of information. Indeed, people who exhibit such qualities appear very intelligent. That said, it seems that intelligence must be more than simply knowing facts and being able to remember them. One point in favor of this argument is the idea of animal intelligence. It will come as no surprise to you that a dog, which can learn commands and tricks seems smarter than a snake that cannot. In fact, researchers and lay people generally agree with one another that primates—monkeys and apes (including humans)—are among the most intelligent animals. Apes such as chimpanzees are capable of complex problem solving and sophisticated communication (Kohler, 1924).

Scientists point to the social nature of primates as one evolutionary source of their intelligence. Primates live together in troops or family groups and are, therefore, highly social creatures. As such, primates tend to have brains that are better developed for communication and long term thinking than most other animals. For instance, the complex social environment has led primates to develop deception, altruism, numerical concepts, and "theory of mind" (a sense of the self as a unique individual separate from others in the group; Gallup, 1982; Hauser, MacNeilage & Ware, 1996).[Also see Noba module Theory of Mind **http://noba.to/a8wpytg3]**

The question of what constitutes human intelligence is one of the oldest inquiries in psychology. When we talk about intelligence we typically mean intellectual ability. This broadly encompasses the ability to learn, remember and use new information, to solve problems and to adapt to novel situations. An early scholar of intelligence, Charles Spearman, proposed the idea that intelligence was one thing, a “general factor” sometimes known as simply “g.” He based this conclusion on the observation that people who perform well in one intellectual area such as verbal ability also tend to perform well in other areas such as logic and reasoning (Spearman, 1904).

A contemporary of Spearman’s named Francis Galton—himself a cousin of Charles Darwin--was among those who pioneered psychological measurement (Hunt, 2009). For three pence Galton would measure various physical characteristics such as grip strength but also some psychological attributes such as the ability to judge distance or discriminate between colors. This is an example of one of the earliest systematic measures of individual ability. Galton was particularly interested in intelligence, which he thought was heritable in much the same way that height and eye color are. He conceived of several rudimentary methods for assessing whether his hypothesis was true. For example, he carefully tracked the family tree of the top-scoring Cambridge students over the previous 40 years. Although he found specific families disproportionately produced top scholars, intellectual achievement could still be the product of economic status, family culture or other non-genetic factors. Galton was also, possibly, the first to popularize the idea that the heritability of psychological traits could be studied by looking at identical and fraternal twins. Although his methods were crude by modern standards, Galton established intelligence as a variable that could be measured (Hunt, 2009).

Intelligence research pioneer Alfred Binet

The person best known for formally pioneering the measurement of intellectual ability is Alfred Binet. Like Galton, Binet was fascinated by individual differences in intelligence. For instance, he blindfolded chess players and saw that some of them had the ability to continue playing using only their memory to keep the many positions of the pieces in mind (Binet, 1894). Binet was particularly interested in the development of intelligence, a fascination that led him to observe children carefully in the classroom setting.

Along with his colleague Theodore Simon, Binet created a test of children’s intellectual capacity. They created individual test items that should be answerable by children of given ages. For instance, a child who is three should be able to point to her mouth and eyes, a child who is nine should be able to name the months of the year in order, and a twelve year old ought to be able to name sixty words in three minutes. Their assessment became the first “IQ test."

1. Which of the following is the most similar to 1313323?

A. ACACCBC

B. CACAABC

C. ABABBCA

D. ACACCDC

2. Jenny has some chocolates. She eats two and gives half of the remainder to Lisa. If Lisa has six chocolates how many does Jenny have in the beginning?

A. 6

B. 12

C. 14

D. 18

3. Which of the following items is not like the others in the list?

duck, raft, canoe, stone, rubber ball

A. Duck

B. Canoe

C. Stone

D. Rubber ball

4. What do steam and ice have in common?

A. They can both harm skin

B. They are both made from water

C. They are both found in the kitchen

D. They are both the products of water at extreme temperatures

Answers: 1) A; 2) C; 3) stone; 4) D is the most sophisticated answer

Table 1: Examples of the types of items you might see on an intelligence test.

"IQ" or "intelligence quotient" is a name given to the score of the Binet-Simon test. The score is derived by dividing a child's mental age (the score from the test) by their chronological age to create an overall quotient. These days, the phrase "IQ" does not apply specifically to the Binet-Simon test and is used to generally denote intelligence or a score on any intelligence test. In the early 1900s the Binet-Simon test was adapted by a Stanford professor named Lewis Terman to create what is, perhaps, the most famous intelligence test in the world, the Stanford-Binet (Terman, 1916). The major advantage of this new test was that it was standardized. Based on a large sample of children Terman was able to plot the scores in a normal distribution, shaped like a "bell curve" (see Fig. 1). To understand a normal distribution think about the height of people. Most people are average in height with relatively fewer being tall or short, and fewer still being extremely tall or extremely short. Terman (1916) laid out intelligence scores in exactly the same way, allowing for easy and reliable categorizations and comparisons between individuals.

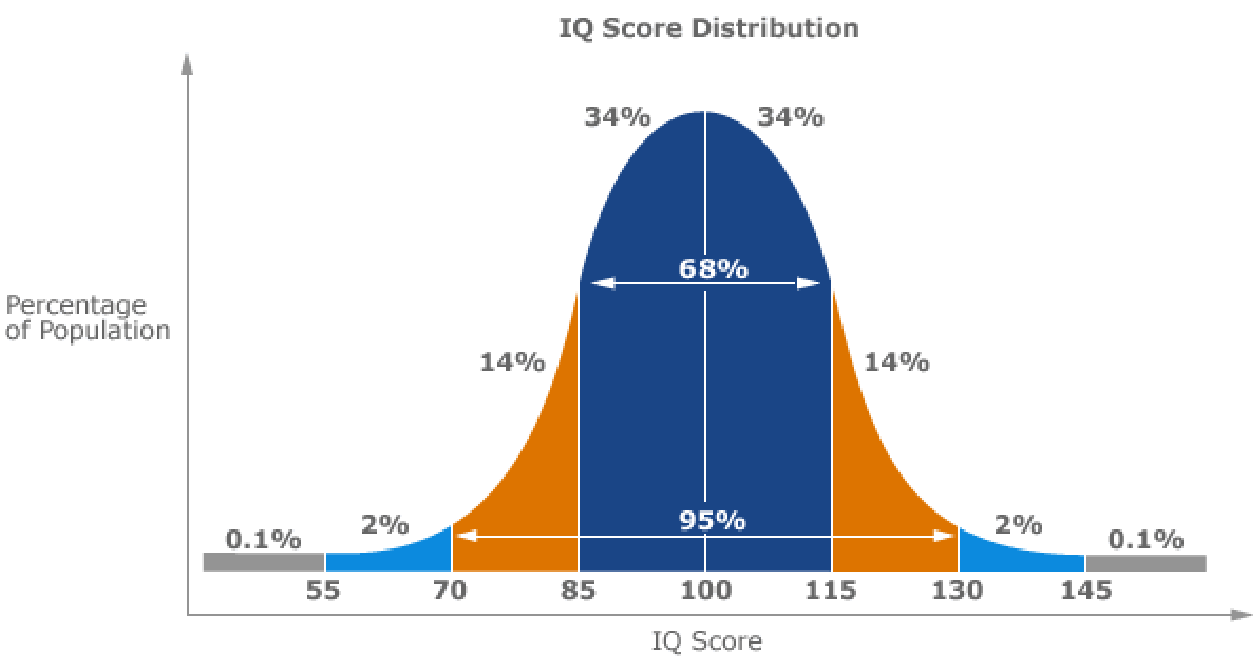


Figure 1: Bell Curve - Normal Distribution IQ

Looking at another modern intelligence test—the Wechsler Adult Intelligence Scale (WAIS)—can provide clues to a definition of intelligence itself. Motivated by several criticisms of the Stanford-Binet test, psychologist David Wechsler sought to create a superior measure of intelligence. He was critical of the way that the Stanford-Binet relied so heavily on verbal ability and was also suspicious of using a single score to capture all of intelligence. To address these issues Wechsler created a test that tapped a wide range of intellectual abilities. This understanding of intelligence—that it is made up of a pool of specific abilities—is a notable departure from Spearman's concept of general intelligence. The WAIS assesses people's ability to remember, compute, understand language, reason well, and process information quickly (Wechsler, 1955).

One interesting by-product of measuring intelligence for so many years is that we can chart changes over time. It might seem strange to you that intelligence can change over the decades but that appears to have happened over the last 80 years we have been measuring this topic. Here's how we know: IQ tests have an average score of 100. When new waves of people are asked to take older tests they tend to outperform the original sample from years ago on which the test was **normed**. This gain is known as the "Flynn Effect," named after James Flynn, the researcher who first identified it (Flynn, 1987). Several hypotheses have been put forth to explain the Flynn Effect including better nutrition (healthier brains!), greater familiarity with testing in general, and more exposure to visual stimuli. Today, there is no perfect agreement among psychological researchers with regards to the causes of increases in average scores on intelligence tests. Perhaps if you choose a career in psychology you will be the one to discover the answer!

Types of Intelligence

David Wechsler's approach to testing intellectual ability was based on the fundamental idea that there are, in essence, many aspects to intelligence. Other scholars have echoed this idea by going so far as to suggest that there are actually even different types of intelligence. You likely have heard distinctions made between "street smarts" and "book learning." The former refers to practical wisdom accumulated through experience while the latter indicates formal education. A person high in street smarts might have a superior ability to catch a person in a lie, to persuade others, or to think quickly under pressure. A person high in book learning, by contrast, might have a large vocabulary and be able to remember a large number of references to classic novels. Although psychologists don't use street smarts or book smarts as professional terms they do believe that intelligence comes in different types.

There are many ways to parse apart the concept of intelligence. Many scholars believe that Carroll 's (1993) review of more than 400 data sets provides the best currently existing single source for organizing various concepts related to intelligence. Carroll divided intelligence into three levels, or strata, descending from the most abstract down to the most specific (see Fig. 2). To understand this way of categorizing simply think of a "car." Car is a general word that denotes all types of motorized vehicles. At the more specific level under "car" might be various types of cars such as sedans, sports cars, SUVs, pick-up trucks, station wagons, and so forth. More specific still would be certain models of each such as a Honda Civic or Ferrari Enzo. In the same manner, Carroll called the highest level (stratum III) the general intelligence factor "g." Under this were more specific stratum II categories such as fluid intelligence and visual perception and processing speed. Each of these, in turn, can be sub-divided into very specific components such as spatial scanning, reaction time, and word fluency.

Thinking of intelligence as Carroll (1993) does, as a collection of specific mental abilities, has helped researchers conceptualize this topic in new ways. For example, Horn and Cattell (1966) distinguish between "fluid" and "crystalized" intelligence, both of which show up on stratum II of Carroll's model. Fluid intelligence is the ability to "think on your feet;" that is, to solve problems. Crystalized intelligence, on the other hand, is the ability to use language, skills and experience to address problems. The former is associated more with youth while the latter increases with age. You may have noticed the way in which younger people can adapt to new situations and use trial and error to quickly figure out solutions. By contrast, older people tend to rely on their relatively superior store of knowledge to solve problems.

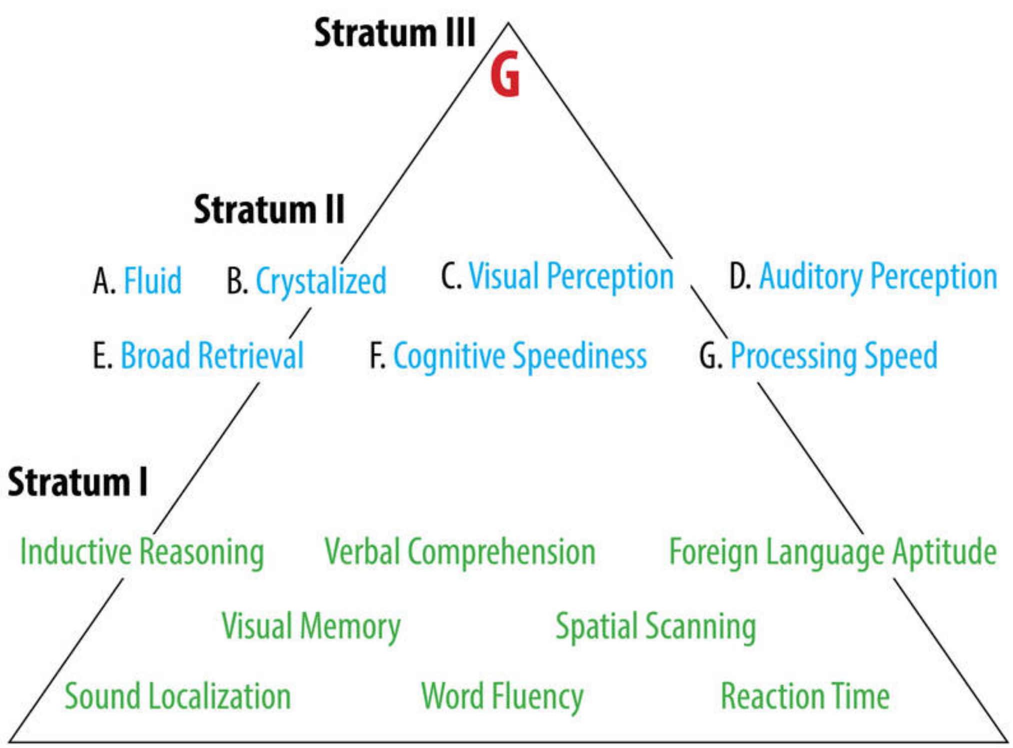


Figure 2: Caroll's Model of Intelligence

Harvard professor Howard Gardner is another figure in psychology who is well-known for championing the notion that there are different types of intelligence. Gardner's theory is appropriately, called "multiple intelligences." Gardner's theory is based on the idea that people process information through different "channels" and these are relatively independent of one another. He has identified 8 common intelligences including 1) logic-math, 2) visual-spatial, 3) music-rhythm, 4) verbal-linguistic, 5) bodily-kinesthetic, 6) interpersonal, 7) intrapersonal, and 8) naturalistic (Gardner, 1985). Many people are attracted to Gardner's theory because it suggests that people each learn in unique ways. There are now many Gardner- influenced schools in the world.

Another type of intelligence is Emotional intelligence. Unlike traditional models of intelligence that emphasize cognition (thinking) the idea of emotional intelligence emphasizes the experience and expression of emotion. Some researchers argue that emotional intelligence is a set of skills in which an individual can accurately understand the emotions of others, can identify and label their own emotions, and can use emotions. (Mayer & Salovey, 1997). Other researchers believe that emotional intelligence is a mixture of abilities, such as stress management, and personality, such as a person's predisposition for certain moods (Bar-On, 2006). Regardless of the specific definition of emotional intelligence, studies have shown a link between this concept and job performance (Lopes, Grewal, Kadis, Gall, & Salovey, 2006). In fact, emotional intelligence is similar to more traditional notions of cognitive intelligence with regards to workplace benefits. Schmidt and Hunter (1998), for example, reviewed research on intelligence in the workplace context and show that intelligence is the single best predictor of doing well in job training programs, of learning on the job. They also report that general intelligence is moderately correlated with all types of jobs but especially with managerial and complex, technical jobs.

There is one last point that is important to bear in mind about intelligence. It turns out that the way an individual thinks about his or her own intelligence is also important because it predicts performance. Researcher Carol Dweck has made a career out of looking at the differences between high IQ children who perform well and those who do not, so-called "under achievers." Among her most interesting findings is that it is not gender or social class that sets apart the high and low performers. Instead, it is their mindset. The children who believe that their abilities in general—and their intelligence specifically—is a fixed trait tend to underperform. By contrast, kids who believe that intelligence is changeable and evolving tend to handle failure better and perform better (Dweck, 1986). Dweck refers to this as a person's "mindset" and having a growth mindset appears to be healthier.

Correlates of Intelligence

The research on mindset is interesting but there can also be a temptation to interpret it as suggesting that every human has an unlimited potential for intelligence and that becoming smarter is only a matter of positive thinking. There is some evidence that genetics is an important factor in the intelligence equation. For instance, a number of studies on genetics in adults have yielded the result that intelligence is largely, but not totally, inherited (Bouchard,2004). Having a healthy attitude about the nature of smarts and working hard can both definitely help intellectual performance but it also helps to have the genetic leaning toward intelligence.

Carol Dweck's research on the mindset of children also brings one of the most interesting and controversial issues surrounding intelligence research to the fore: group differences. From the very beginning of the study of intelligence researchers have wondered about differences between groups of people such as men and women. With regards to potential differences between the sexes some people have noticed that women are under-represented in certain fields. In 1976, for example, women comprised just 1% of all faculty members in engineering (Ceci, Williams & Barnett, 2009).

Women account for a disproportionately small percentage of those employed in math-intensive career fields such as engineering. [Photo: Argonne National Laboratory, https://goo.gl/ix96YP, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Even today women make up between 3% and 15% of all faculty in math-intensive fields at the 50 top universities. This phenomenon could be explained in many ways: it might be the result of inequalities in the educational system, it might be due to differences in socialization wherein young girls are encouraged to develop other interests, it might be the result of that women are—on average—responsible for a larger portion of childcare obligations and therefore make different types of professional decisions, or it might be due to innate differences between these groups, to name just a few possibilities. The possibility of innate differences is the most controversial because many people see it as either the product of or the foundation for sexism. In today's political landscape it is easy to see that asking certain questions such as "are men smarter than women?" would be inflammatory. In a comprehensive review of research on intellectual abilities and sex Ceci and colleagues (2009) argue against the hypothesis that biological and genetic differences account for much of the sex differences in intellectual ability. Instead, they believe that a complex web of influences ranging from societal expectations to test taking strategies to individual interests account for many of the sex differences found in math and similar intellectual abilities.

A more interesting question, and perhaps a more sensitive one, might be to inquire in which ways men and women might differ in intellectual ability, if at all. That is, researchers should not seek to prove that one group or another is better but might examine the ways that they might differ and offer explanations for any differences that are found. Researchers have investigated sex differences in intellectual ability. In a review of the research literature Halpern (1997) found that women appear, on average, superior to men on measures of fine motor skill, acquired knowledge, reading comprehension, decoding non-verbal expression, and generally have higher grades in school. Men, by contrast, appear, on average, superior to women on measures of fluid reasoning related to math and science, perceptual tasks that involve moving objects, and tasks that require transformations in working memory such as mental rotations of physical spaces. Halpern also notes that men are disproportionately represented on the low end of cognitive functioning including in intellectual disability , dyslexia, and attention deficit disorders (Halpern, 1997).

Other researchers have examined various explanatory hypotheses for why sex differences in intellectual ability occur. Some studies have provided mixed evidence for genetic factors while others point to evidence for social factors (Neisser, et al, 1996; Nisbett, et al., 2012). One interesting phenomenon that has received research scrutiny is the idea of **stereotype threat**. Stereotype threat is the idea that mental access to a particular stereotype can have real-world impact on a member of the stereotyped group. In one study (Spencer, Steele, & Quinn, 1999), for example, women who were informed that women tend to fare poorly on math exams just before taking a math test actually performed worse relative to a control group who did not hear the stereotype. Research on stereotype has yielded mixed results and we are currently uncertain about exactly how and when this effect might occur. One possible antidote to stereotype threat, at least in the case of women, is to make a self-affirmation (such as listing positive personal qualities) before the threat occurs. In one study, for instance, Martens and her colleagues (2006) had women write about personal qualities that they valued before taking a math test. The affirmation largely erased the effect of stereotype by improving math scores for women relative to a control group but similar affirmations had little effect for men (Martens, Johns, Greenberg, & Schimel, 2006).

These types of controversies compel many lay people to wonder if there might be a problem with intelligence measures. It is natural to wonder if they are somehow biased against certain groups. Psychologists typically answer such questions by pointing out that bias in the testing sense of the word is different than how people use the word in everyday speech. Common use of bias denotes a prejudice based on group membership. Scientific bias, on the other hand, is related to the psychometric properties of the test such as validity and reliability.

Validity is the idea that an assessment measures what it claims to measure and that it can predict future behaviors or performance. To this end, intelligence tests are not biased because they are fairly accurate measures and predictors. There are, however, real biases, prejudices, and inequalities in the social world that might benefit some advantaged group while hindering some disadvantaged others.

Conclusion

Although you might not be able to spell “esquamulose” or “staphylococci” – indeed, you might not even know what they mean—you don’t need to count yourself out in the intelligence department. Now that we have examined intelligence in depth we can return to our intuitive view of those students who compete in the National Spelling Bee. Are they smart? Certainly, they seem to have high verbal intelligence. There is also the possibility that they benefit from either a genetic boost in intelligence, a supportive social environment, or both. Watching them spell difficult words there is also much we do not know about them. We cannot tell, for instance, how emotionally intelligent they are or how they might use bodily-kinesthetic intelligence. This highlights the fact that intelligence is a complicated issue. Fortunately, psychologists continue to research this fascinating topic and their studies continue to yield new insights.

Outside Resources

Blog: Dr. Jonathan Wai has an excellent blog on Psychology Today discussing many of the most interesting issue related to intelligence.
http://www.psychologytoday.com/blog/finding-the-next-einstein

Video: Hank Green gives a fun and interesting overview of the concept of intelligence in this installment of the Crash Course series.
https://www.youtube.com/watch?v=9xTz3QjcloI

Discussion Questions

1. Do you think that people get smarter as they get older? In what ways might people gain or lose intellectual abilities as they age?
2. When you meet someone who strikes you as being smart what types of cues or information do you typically attend to in order to arrive at this judgment?
3. How do you think socio-economic status affects an individual taking an intellectual abilities test?
4. Should psychologists be asking about group differences in intellectual ability? What do you think?
5. Which of Howard Gardner's 8 types of intelligence do you think describes the way you learn best?

Vocabulary

G
Short for "general factor" and is often used to be synonymous with intelligence itself.

Intelligence
An individual's cognitive capability. This includes the ability to acquire, process, recall and apply information.

IQ
Short for "intelligence quotient." This is a score, typically obtained from a widely used measure of intelligence that is meant to rank a person's intellectual ability against that of others.

Norm
Assessments are given to a representative sample of a population to determine the range of scores for that population. These "norms" are then used to place an individual who takes that assessment on a range of scores in which he or she is compared to the population at large.

Standardize
Assessments that are given in the exact same manner to all people . With regards to intelligence tests standardized scores are individual scores that are computed to be referenced against normative scores for a population (see "norm").

Stereotype threat
The phenomenon in which people are concerned that they will conform to a stereotype or that their performance does conform to that stereotype, especially in instances in which the stereotype is brought to their conscious awareness.

References

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicometha, 18* (Suppl.), 13–25.

Binet, A. (1894). *Psychologie des grands calculateurs et joueurs d'échecs*. Paris: Librairie Hachette.

Bouchard, T.J. (2004). Genetic influence on human psychological traits - A survey. *Current Directions in Psychological Science 13*(4), 148–151.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England:Cambridge University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England:Cambridge University Press.

Ceci, S. J., Williams, W. & Barnett, S. M. (2009). Women's underrepresentation in science: socio cultural and biological considerations. *Psychological Bulletin, 135*, 218-261.

Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist, 41*(10), 1040-1048.

Flynn J. R. (1987). "Massive IQ gains in 14 nations: What IQ tests really measure". *Psychological Bulletin 101*, 171–191.

Gallup, G. G. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology, 2*(3), 237-248.

Gardner, H. (1985). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist, 52*(10), 1091-1102.

Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist, 52*(10), 1091-1102.

Hauser, M. D., MacNeilage, P., & Ware, M. (1996). Numerical representations in primates. *Proceedings of the National Academy of Sciences, 93*(4), 1514-1517.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*(5), 253-270.

Hunt, M. (2009). *The story of psychology*. New York: Random House, LLC.

Hunt, M. (2009). *The story of psychology*. New York: Random House, LLC.

Kohler, W. (1924). *The mentality of apes*. Oxford: Harcourt, Brace.

Lopes, P. N., Grewal, D., Kadis, J., Gall, M., & Salovey, P. (2006). Evidence that emotional intelligence is related to job performance and affect and attitudes at work. *Psicothema, 18*

(Suppl.), 132–138.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women’s intellectual performance. *Journal of Experimental Social Psychology, 42*(2), 236-243.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women’s intellectual performance. *Journal of Experimental Social Psychology, 42*(2), 236-243.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), *Emotional development and emotional intelligence: Educational implications* (pp. 3–34). New York: Basic.

National Spelling Bee. (2014a). *Statistics*. Retrieved from: http://www.spellingbee.com/statistics

National Spelling Bee. (2014b). *Get to Know the Competition*. Retrieved from: http://www.spellingbee.com/UserFiles/topblog----good2341418.html

Neisser, U., Boodoo, G., Bouchard, Jr., T.J., Boykin, A. W., Brody, N., Ceci, S., Halpern, D., Loehlin, J. C., Perloff, R., Sternberg, R. J. & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *American Psychologist, 67*(2), 130-160.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–274

Spearman, C. (1904). " General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology, 15*(2), 201-292.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale*. Boston: Houghton Mifflin.

Terman, L. M. (1916). *The measurement of intelligence: An explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon Intelligence Scale*. Boston: Houghton Mifflin.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. Oxford: Psychological Corporation.

25
Judgment and Decision Making

Max H. Bazerman

Humans are not perfect decision makers. Not only are we not perfect, but we depart from perfection or rationality in systematic and predictable ways. The understanding of these systematic and predictable departures is core to the field of judgment and decision making. By understanding these limitations, we can also identify strategies for making better and more effective decisions.

Learning Objectives

- Understand the systematic biases that affect our judgment and decision making.
- Develop strategies for making better decisions.
- Experience some of the biases through sample decisions.

Introduction

Every day you have the opportunity to make countless decisions: should you eat dessert, cheat on a test, or attend a sports event with your friends. If you reflect on your own history of choices you will realize that they vary in quality; some are rational and some are not. This module provides an overview of decision making and includes discussion of many of the common biases involved in this process.

In his Nobel Prize–winning work, psychologist Herbert Simon (1957; March & Simon, 1958) argued that our decisions are bounded in their rationality. According to the **bounded** rationality framework, human beings try to make rational decisions (such as weighing the costs and benefits of a choice) but our cognitive limitations prevent us from being fully rational. Time and cost constraints limit the quantity and quality of the information that is available to us. Moreover, we only retain a relatively small amount of information in our usable memory. And limitations on intelligence and perceptions constrain the ability of even very bright decision makers to accurately make the best choice based on the information that is available.

About 15 years after the publication of Simon's seminal work, Tversky and Kahneman (1973, 1974; Kahneman & Tversky, 1979) produced their own Nobel Prize–winning research, which provided critical information about specific systematic and predictable biases, or mistakes, that influence judgment (Kahneman received the prize after Tversky's death). The work of Simon, Tversky, and Kahneman paved the way to our modern understanding of judgment and decision making. And their two Nobel prizes signaled the broad acceptance of the field of behavioral decision research as a mature area of intellectual study.

What Would a Rational Decision Look Like?

People often have to use incomplete information and intuition to make even the most important of decisions. A fully rational decision requires a careful, systematic process. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Imagine that during your senior year in college, you apply to a number of doctoral programs, law schools, or business schools (or another set of programs in whatever field most interests you). The good news is that you receive many acceptance letters. So, how should you decide where to go? Bazerman and Moore (2013) outline the following six steps that you should take to make a rational decision: (1) define the problem (i.e., selecting the right graduate program), (2) identify the criteria necessary to judge the multiple options (location, prestige, faculty, etc.), (3) weight the criteria (rank them in terms of importance to you), (4) generate alternatives (the schools that admitted you), (5) rate each alternative on each criterion (rate each school on each criteria that you identified, and (6) compute the optimal decision. Acting rationally would require that you follow these six steps in a fully rational manner.

I strongly advise people to think through important decisions such as this in a manner similar to this process. Unfortunately, we often don't. Many of us rely on our intuitions far more than we should. And when we do try to think systematically, the way we enter data into such formal decision-making processes is often biased.

Fortunately, psychologists have learned a great deal about the biases that affect our thinking. This knowledge about the systematic and predictable mistakes that even the best and the brightest make can help you identify flaws in your thought processes and reach better decisions.

Biases in Our Decision Process

Simon's concept of bounded rationality taught us that judgment deviates from rationality, but it did not tell us *how* judgment is biased. Tversky and Kahneman's (1974) research helped to diagnose the specific systematic, directional biases that affect human judgment. These biases are created by the tendency to short-circuit a rational decision process by relying on a number of simplifying strategies, or rules of thumb, known as **heuristics**. Heuristics allow us to cope with the complex environment surrounding our decisions. Unfortunately, they also lead to systematic and predictable biases.

To highlight some of these biases please answer the following three quiz items:

Problem 1 (adapted from Alpert & Raiffa, 1969):

Listed below are 10 uncertain quantities. Do not look up any information on these items. For each, write down your best estimate of the quantity. Next, put a lower and upper bound around your estimate, such that you are 98 percent confident that your range surrounds the actual quantity. Respond to each of these items even if you admit to knowing very little about these quantities.

1. The first year the Nobel Peace Prize was awarded
2. The date the French celebrate "Bastille Day"
3. The distance from the Earth to the Moon
4. The height of the Leaning Tower of Pisa
5. Number of students attending Oxford University (as of 2014)
6. Number of people who have traveled to space (as of 2013)

7. 2012-2013 annual budget for the University of Pennsylvania
8. Average life expectancy in Bangladesh (as of 2012)
9. World record for pull-ups in a 24-hour period
10. Number of colleges and universities in the Boston metropolitan area

Problem 2 (adapted from Joyce & Biddle, 1981):

We know that executive fraud occurs and that it has been associated with many recent financial scandals. And, we know that many cases of management fraud go undetected even when annual audits are performed. Do you think that the incidence of significant executive-level management fraud is more than 10 in 1,000 firms (that is, 1 percent) audited by Big Four accounting firms?

1. Yes, more than 10 in 1,000 Big Four clients have significant executive-level management fraud.
2. No, fewer than 10 in 1,000 Big Four clients have significant executive-level management fraud.

What is your estimate of the number of Big Four clients per 1,000 that have significant executive-level management fraud? (Fill in the blank below with the appropriate number.)

________ in 1,000 Big Four clients have significant executive-level management fraud.

Problem 3 (adapted from Tversky & Kahneman, 1981):

Imagine that the United States is preparing for the outbreak of an unusual avian disease that is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows.

1. Program A: If Program A is adopted, 200 people will be saved.
2. Program B: If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.

Which of the two programs would you favor?

Overconfidence

On the first problem, if you set your ranges so that you were justifiably 98 percent confident, you should expect that approximately 9.8, or nine to 10, of your ranges would include the actual value. So, let's look at the correct answers:

1. 1901
2. 14th of July
3. 384,403 km (238,857 mi)
4. 56.67 m (183 ft)
5. 22,384 (as of 2014)
6. 536 people (as of 2013)
7. $6.007 billion
8. 70.3 years (as of 2012)
9. 4,321
10. 52

Overconfidence is a natural part of most people's decision-making process and this can get us into trouble. Is it possible to overcome our faulty thinking? Perhaps. See the "Fixing Our Decisions" section below. [Image: Barn Images, https://goo.gl/IYzbDV, CC BY 2.0, https://goo.gl/BRvSA7]

Count the number of your 98% ranges that actually surrounded the true quantities. If you surrounded nine to 10, you were appropriately confident in your judgments. But most readers surround only between three (30%) and seven (70%) of the correct answers, despite claiming 98% confidence that each range would surround the true value. As this problem shows, humans tend to be overconfident in their judgments.

Anchoring

Regarding the second problem, people vary a great deal in their final assessment of the level of executive-level management fraud, but most think that 10 out of 1,000 is too low. When I run this exercise in class, half of the students respond to the question that I asked you to answer. The other half receive a similar problem, but instead are asked whether the correct answer is higher or lower than 200 rather than 10. Most people think that 200 is high. But, again, most people claim that this "anchor" does not affect their final estimate. Yet, on average, people who are presented with the question that focuses on the number 10 (out of 1,000) give answers that are about one-half the size of the estimates of those facing questions that use an anchor of 200. When we are making decisions, any initial anchor that we face is likely to influence our judgments, even if the anchor is arbitrary. That is, we insufficiently adjust our judgments away from the anchor.

Framing

Turning to Problem 3, most people choose Program A, which saves 200 lives for sure, over Program B. But, again, if I was in front of a classroom, only half of my students would receive this problem. The other half would have received the same set-up, but with the following two options:

1. Program C: If Program C is adopted, 400 people will die.
2. Program D: If Program D is adopted, there is a one-third probability that no one will die and a two-thirds probability that 600 people will die.

Which of the two programs would you favor?

Careful review of the two versions of this problem clarifies that they are objectively the same. Saving 200 people (Program A) means losing 400 people (Program C), and Programs B and D are also objectively identical. Yet, in one of the most famous problems in judgment and decision making, most individuals choose Program A in the first set and Program D in the second set (Tversky & Kahneman, 1981). People respond very differently to saving versus losing lives—even when the difference is based just on the "**framing**" of the choices.

The problem that I asked you to respond to was framed in terms of saving lives, and the implied reference point was the worst outcome of 600 deaths. Most of us, when we make decisions that concern gains, are risk averse; as a consequence, we lock in the possibility of saving 200 lives for sure. In the alternative version, the problem is framed in terms of losses. Now the implicit reference point is the best outcome of no deaths due to the avian disease. And in this case, most people are risk seeking when making decisions regarding losses.

These are just three of the many biases that affect even the smartest among us. Other research shows that we are biased in favor of information that is easy for our minds to retrieve, are insensitive to the importance of base rates and sample sizes when we are making inferences, assume that random events will always look random, search for information that confirms our expectations even when disconfirming information would be more informative, claim a priori knowledge that didn't exist due to the hindsight bias, and are subject to a host of other effects that continue to be developed in the literature (Bazerman & Moore, 2013).

Contemporary Developments

The concept of bounded willpower may explain why many of us are better shoppers than savers. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Bounded rationality served as the integrating concept of the field of behavioral decision research for 40 years. Then, in 2000, Thaler (2000) suggested that decision making is bounded in two ways not precisely captured by the concept of bounded rationality. First, he argued that our **willpower is bounded** and that, as a consequence, we give greater weight to present concerns than to future concerns. Our immediate motivations are often inconsistent with our long-term interests in a variety of ways, such as the common failure to save adequately for retirement or the difficulty many people have staying on a diet. Second, Thaler suggested that our **self-interest is bounded** such that we care about the outcomes of others. Sometimes we positively value the outcomes of others—giving them more of a commodity than is necessary out of a desire to be fair, for example. And, in unfortunate contexts, we sometimes are willing to forgo our own benefits out of a desire to harm others.

My colleagues and I have recently added two other important bounds to the list. Chugh, Banaji, and Bazerman (2005) and Banaji and Bhaskar (2000) introduced the concept of **bounded ethicality**, which refers to the notion that our ethics are limited in ways we are not even aware of ourselves. Second, Chugh and Bazerman (2007) developed the concept of **bounded awareness** to refer to the broad array of focusing failures that affect our judgment, specifically the many ways in which we fail to notice obvious and important information that is available to us.

A final development is the application of judgment and decision-making research to the areas of behavioral economics, behavioral finance, and behavioral marketing, among others. In each case, these fields have been transformed by applying and extending research from the judgment and decision-making literature.

Fixing Our Decisions

Ample evidence documents that even smart people are routinely impaired by biases. Early research demonstrated, unfortunately, that awareness of these problems does little to reduce bias (Fischhoff, 1982). The good news is that more recent research documents interventions that do help us overcome our faulty thinking (Bazerman & Moore, 2013).

One critical path to fixing our biases is provided in Stanovich and West's (2000) distinction between **System 1** and **System 2** decision making. System 1 processing is our intuitive system, which is typically fast, automatic, effortless, implicit, and emotional. System 2 refers to decision making that is slower, conscious, effortful, explicit, and logical. The six logical steps of decision making outlined earlier describe a System 2 process.

Clearly, a complete System 2 process is not required for every decision we make. In most situations, our System 1 thinking is quite sufficient; it would be impractical, for example, to logically reason through every choice we make while shopping for groceries. But, preferably, System 2 logic should influence our most important decisions. Nonetheless, we use our System 1 processes for most decisions in life, relying on it even when making important decisions.

The key to reducing the effects of bias and improving our decisions is to transition from trusting our intuitive System 1 thinking toward engaging more in deliberative System 2 thought. Unfortunately, the busier and more rushed people are, the more they have on their minds, and the more likely they are to rely on System 1 thinking (Chugh, 2004). The frantic pace of professional life suggests that executives often rely on System 1 thinking (Chugh, 2004).

Fortunately, it is possible to identify conditions where we rely on intuition at our peril and substitute more deliberative thought. One fascinating example of this substitution comes from journalist Michael Lewis' (2003) account of how Billy Beane, the general manager of the Oakland Athletics, improved the outcomes of the failing baseball team after recognizing that the intuition of baseball executives was limited and systematically biased and that their intuitions had been incorporated into important decisions in ways that created enormous mistakes. Lewis (2003) documents that baseball professionals tend to overgeneralize from their personal experiences, be overly influenced by players' very recent performances, and overweigh what they see with their own eyes, despite the fact that players' multiyear records provide far better data. By substituting valid predictors of future performance (System 2 thinking), the Athletics were able to outperform expectations given their very limited payroll.

Nudges can be used to help people make better decisions about saving for retirement. [Image: Tax Credits, https://goo.gl/YLuyth, CC BY 2.0, https://goo.gl/BRvSA7]

Another important direction for improving decisions comes from Thaler and Sunstein's (2008) book *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Rather than setting out to debias human judgment, Thaler and Sunstein outline a strategy for how "decision architects" can change environments in ways that account for human bias and trigger better decisions as a result. For example, Beshears, Choi, Laibson, and Madrian (2008) have shown that simple changes to defaults can dramatically improve people's decisions. They tackle the failure of many people to save for retirement and show that a simple change can significantly influence enrollment in 401(k) programs. In most companies, when you start your job, you need to proactively sign up to join the company's retirement savings plan. Many people take years before getting around to doing so. When, instead, companies automatically enroll their employees in 401(k) programs and give them the opportunity to "opt out," the net enrollment rate rises significantly. By changing defaults, we can counteract the human tendency to live with the status quo.

Similarly, Johnson and Goldstein's (2003) cross-European organ donation study reveals that countries that have opt-in organ donation policies, where the default is not to harvest people's organs without their prior consent, sacrifice thousands of lives in comparison to opt-out policies, where the default is to harvest organs. The United States and too many other countries require that citizens opt in to organ donation through a proactive effort; as a consequence, consent rates range between 4.25%–44% across these countries. In contrast, changing the decision architecture to an opt-out policy improves consent rates to 85.9% to 99.98%. Designing the donation system with knowledge of the power of defaults can dramatically change donation rates without changing the options available to citizens. In contrast, a more intuitive strategy, such as the one in place in the United States, inspires defaults that result in many unnecessary deaths.

Concluding Thoughts

Our days are filled with decisions ranging from the small (what should I wear today?) to the important (should we get married?). Many have real world consequences on our health, finances and relationships. Simon, Kahneman, and Tversky created a field that highlights the surprising and predictable deficiencies of the human mind when making decisions. As we understand more about our own biases and thinking shortcomings we can begin to take them into account or to avoid them. Only now have we reached the frontier of using this knowledge to help people make better decisions.

Outside Resources

Book: Bazerman, M. H., & Moore, D. (2013). Judgment in managerial decision making (8th ed.). John Wiley & Sons Inc.

Book: Kahneman, D. (2011) Thinking, Fast and Slow. New York, NY: Farrar, Straus and Giroux.

Book: Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving Decisions about Health, Wealth, and Happiness. New Haven, CT: Yale University Press.

Discussion Questions

1. Are the biases in this module a problem in the real world?
2. How would you use this module to be a better decision maker?
3. Can you see any biases in today's newspaper?

Vocabulary

Anchoring
The bias to be affected by an initial anchor, even if the anchor is arbitrary, and to insufficiently adjust our judgments away from that anchor.

Biases
The systematic and predictable mistakes that influence the judgment of even very talented human beings.

Bounded awareness
The systematic ways in which we fail to notice obvious and important information that is available to us.

Bounded ethicality
The systematic ways in which our ethics are limited in ways we are not even aware of ourselves.

Bounded rationality
Model of human behavior that suggests that humans try to make rational decisions but are bounded due to cognitive limitations.

Bounded self-interest
The systematic and predictable ways in which we care about the outcomes of others.

Bounded willpower
The tendency to place greater weight on present concerns rather than future concerns.

Framing
The bias to be systematically affected by the way in which information is presented, while holding the objective information constant.

Heuristics
cognitive (or thinking) strategies that simplify decision making by using mental short-cuts

Overconfident
The bias to have greater confidence in your judgment than is warranted based on a rational assessment.

System 1
Our intuitive decision-making system, which is typically fast, automatic, effortless, implicit, and emotional.

System 2
Our more deliberative decision-making system, which is slower, conscious, effortful, explicit, and logical.

References

Alpert, M., & Raiffa, H. (1969). A progress report on the training of probability assessors. Unpublished Report.

Banaji, M. R., & Bhaskar, R. (2000). Implicit stereotypes and memory: The bounded rationality of social beliefs. In D. L. Schacter & E. Scarry (Eds.), *Memory, brain, and belief* (pp. 139–175). Cambridge, MA: Harvard University Press.

Bazerman, M. H., & Moore, D. (2013). *Judgment in managerial decision making* (8th ed.). John Wiley & Sons Inc.

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2008). The importance of default options for retirement saving outcomes: Evidence from the United States. In S. J. Kay & T. Sinha (Eds.), *Lessons from pension reform in the Americas* (pp. 59–87). Oxford: Oxford University Press.

Chugh, D. (2004). Societal and managerial implications of implicit social cognition: Why milliseconds matter. *Social Justice Research*, 17(2), 203–222.

Chugh, D., & Bazerman, M. H. (2007). Bounded awareness: What you fail to see can hurt you. *Mind & Society*, 6(1), 1–18.

Chugh, D., Banaji, M. R., & Bazerman, M. H. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. In D. Moore, D. M. Cain, G. Loewenstein, & M. H. Bazerman (Eds.), *Conflicts of Interest* (pp. 74–95). New York, NY: Cambridge University Press.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). New York, NY: Cambridge University Press.

Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science 302*(5649), 1338–1339.

Joyce, E. J., & Biddle, G. C. (1981). Are auditors' judgments sufficiently regressive? *Journal of Accounting Research*, 19(2), 323–349.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.

Lewis, M. (2003). *Moneyball: The art of winning an unfair game*. New York, NY: W.W. Norton & Company Ltd.

March, J. G., & Simon, H. A. (1958). *Organizations*. Oxford: Wiley.

Simon, H. A. (1957). *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*. New York, NY: John Wiley & Sons.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645–726.

Thaler, R. H. (2000). From homo economicus to homo sapiens. *Journal of Economics Perspectives*, 14, 133–141.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* New Haven, CT: Yale University Press.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, New Series, 211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, New Series, 185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207–232.

Development

26
Cognitive Development in Childhood

Robert Siegler

This module examines what cognitive development is, major theories about how it occurs, the roles of nature and nurture, whether it is continuous or discontinuous, and how research in the area is being used to improve education.

Learning Objectives

- Be able to identify and describe the main areas of cognitive development.
- Be able to describe major theories of cognitive development and what distinguishes them.
- Understand how nature and nurture work together to produce cognitive development.
- Understand why cognitive development is sometimes viewed as discontinuous and sometimes as continuous.
- Know some ways in which research on cognitive development is being used to improve education.

Introduction

By the time you reach adulthood you have learned a few things about how the world works. You know, for instance, that you can't walk through walls or leap into the tops of trees. You know that although you cannot see your car keys they've got to be around here someplace. What's more, you know that if you want to communicate complex ideas like ordering a triple-shot soy vanilla latte with chocolate sprinkles it's better to use words with meanings attached to them rather than simply gesturing and grunting. People accumulate all this useful knowledge through the process of cognitive development, which involves a multitude of factors, both inherent and learned.

Cognitive development refers to the development of thinking across the lifespan. Defining thinking can be problematic, because no clear boundaries separate thinking from other mental activities. Thinking obviously involves the higher mental processes: problem solving, reasoning, creating, conceptualizing, categorizing, remembering, planning, and so on. However, thinking also involves other mental processes that seem more basic and at which even toddlers are skilled—such as perceiving objects and events in the environment, acting skillfully on objects to obtain goals, and understanding and producing language. Yet other areas of human development that involve thinking are not usually associated with cognitive development, because thinking isn't a prominent feature of them—such as personality and temperament.

Cognitive development in childhood is about change. From birth to adolescence a young person's mind changes dramatically in many important ways. [Image: One Laptop per Child, https://goo.gl/L1eAsO, CC BY 2.0, https://goo.gl/9uSnqN]

As the name suggests, cognitive development is about change. Children's thinking changes in dramatic and surprising ways. Consider DeVries's (1969) study of whether young children understand the difference between appearance and reality. To find out, she brought an unusually even-tempered cat named Maynard to a psychology laboratory and allowed the 3- to 6-year-old participants in the study to pet and play with him. DeVries then put a mask of a fierce dog on Maynard's head, and asked the children what Maynard was. Despite all of the children having identified Maynard previously as a cat, now most 3-year-olds said that he was a dog and claimed that he had a dog's bones and a dog's stomach. In contrast, the 6-year-olds weren't fooled; they had no doubt that Maynard remained a cat. Understanding how children's thinking changes so dramatically in just a few years is one of the fascinating challenges in studying cognitive development.

There are several main types of theories of child development. Stage theories, such as Piaget's stage theory, focus on whether children progress through qualitatively different stages of development. **Sociocultural theories**, such as that of Lev Vygotsky, emphasize how other people and the attitudes, values, and beliefs of the surrounding culture, influence children's development. **Information processing theories**, such as that of David Klahr, examine the mental processes that produce thinking at any one time and the transition processes that lead to growth in that thinking.

At the heart of all of these theories, and indeed of all research on cognitive development, are two main questions: (1) How do nature and nurture interact to produce cognitive development? (2) Does cognitive development progress through qualitatively distinct stages? In the remainder of this module, we examine the answers that are emerging regarding these questions, as well as ways in which cognitive developmental research is being used to improve education.

Nature and Nurture

The most basic question about child development is how nature and nurture together shape development. **Nature** refers to our biological endowment, the genes we receive from our parents. **Nurture** refers to the environments, social as well as physical, that influence our development, everything from the womb in which we develop before birth to the homes in which we grow up, the schools we attend, and the many people with whom we interact.

The nature-nurture issue is often presented as an either-or question: Is our intelligence (for example) due to our genes or to the environments in which we live? In fact, however, every aspect of development is produced by the interaction of genes and environment. At the most basic level, without genes, there would be no child, and without an environment to provide nurture, there also would be no child.

The way in which nature and nurture work together can be seen in findings on visual development. Many people view vision as something that people either are born with or that is purely a matter of biological maturation, but it also depends on the right kind of experience at the right time. For example, development of **depth perception**, the ability to actively perceive the distance from oneself to objects in the environment, depends on seeing patterned light and having normal brain activity in response to the patterned light, in infancy (Held, 1993). If no patterned light is received, for example when a baby has severe cataracts or blindness that is not surgically corrected until later in development, depth perception remains abnormal even after the surgery.

Adding to the complexity of the nature-nurture interaction, children's genes lead to their

A child that is perceived to be attractive and calm may receive a different sort of care and attention from adults and as a result enjoy a developmental advantage. [Image: Cairn 111, https://goo.gl/6RpBVt, CC BY-NC-SA 2.0, https://goo.gl/HEXbAA]

eliciting different treatment from other people, which influences their cognitive development. For example, infants' physical attractiveness and temperament are influenced considerably by their genetic inheritance, but it is also the case that parents provide more sensitive and affectionate care to easygoing and attractive infants than to difficult and less attractive ones, which can contribute to the infants' later cognitive development (Langlois et al., 1995; van den Boom & Hoeksma, 1994).

Also contributing to the complex interplay of nature and nurture is the role of children in shaping their own cognitive development. From the first days out of the womb, children actively choose to attend more to some things and less to others. For example, even 1-month-olds choose to look at their mother's face more than at the faces of other women of the same age and general level of attractiveness (Bartrip, Morton, & de Schonen, 2001). Children's contributions to their own cognitive development grow larger as they grow older (Scarr & McCartney, 1983). When children are young, their parents largely determine their experiences: whether they will attend day care, the children with whom they will have play dates, the books to which they have access, and so on. In contrast, older children and adolescents choose their environments to a larger degree. Their parents' preferences largely determine how 5-year-olds spend time, but 15-year-olds' own preferences largely determine when, if ever, they set foot in a library. Children's choices often have large consequences. To cite one example, the more that children choose to read, the more that their reading improves in future years (Baker, Dreher, & Guthrie, 2000). Thus, the issue is not whether cognitive development is a product of nature or nurture; rather, the issue is how nature and nurture work together to produce cognitive development.

Does Cognitive Development Progress Through Distinct Stages?

Some aspects of the development of living organisms, such as the growth of the width of a pine tree, involve quantitative changes, with the tree getting a little wider each year. Other changes, such as the life cycle of a ladybug, involve qualitative changes, with the creature becoming a totally different type of entity after a transition than before (Figure 1). The existence of both gradual, quantitative changes and relatively sudden, qualitative changes in the world has led researchers who study cognitive development to ask whether changes in children's thinking are gradual and **continuous** or sudden and **discontinuous.**

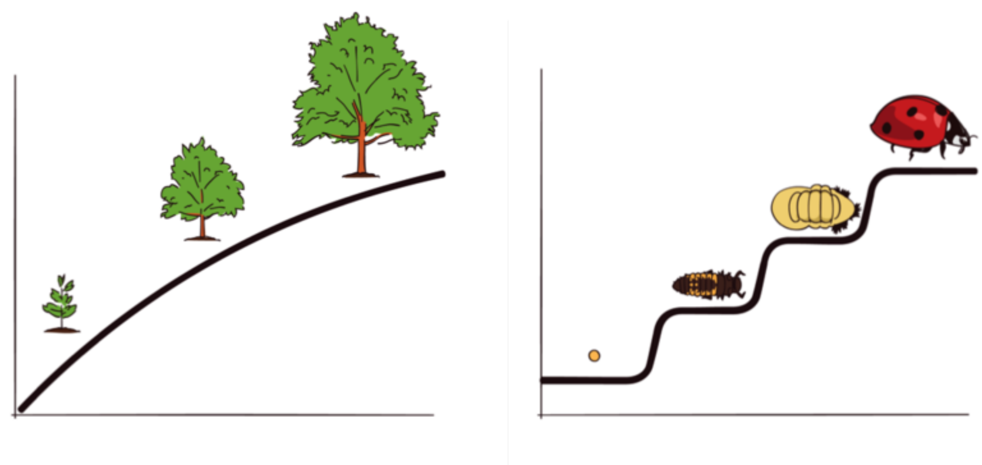
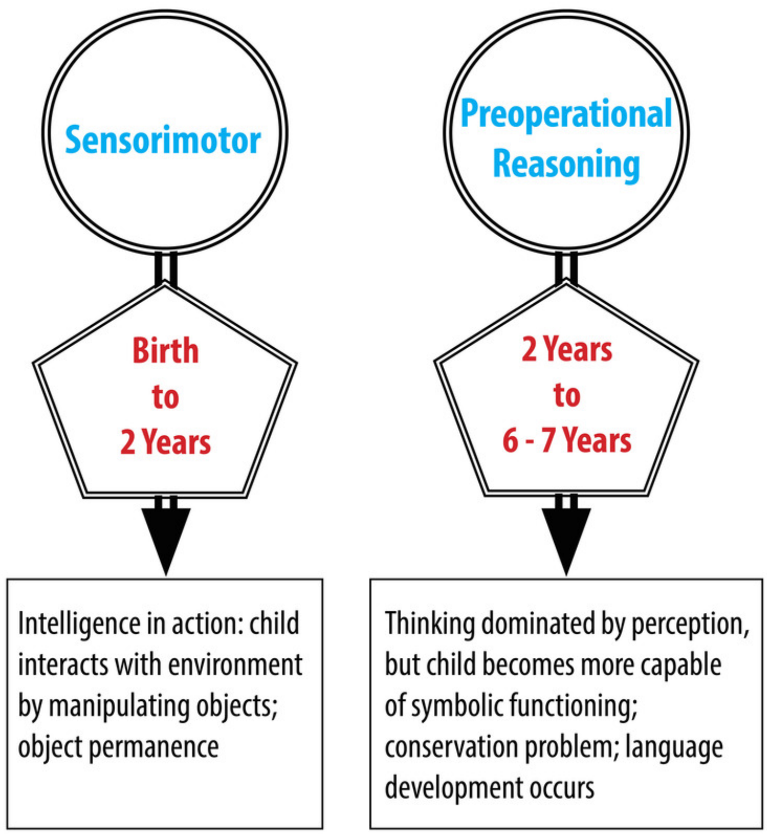


Figure 1: Continuous and discontinuous development. Some researchers see development as a continuous gradual process, much like a maple tree growing steadily in height and cross-sectional area. Other researchers see development as a progression of discontinuous stages, involving rapid discontinuous changes, such as those in the life cycle of a ladybug, separated by longer periods of slow, gradual change.

The great Swiss psychologist Jean Piaget proposed that children's thinking progresses through a series of four discrete stages. By "stages," he meant periods during which children reasoned similarly about many superficially different problems, with the stages occurring in a fixed order and the thinking within different stages differing in fundamental ways. The four stages that Piaget hypothesized were the **sensorimotor stage** (birth to 2 years), the **preoperational reasoning stage** (2 to 6 or 7 years), the **concrete operational reasoning stage** (6 or 7 to 11 or 12 years), and the **formal operational reasoning stage** (11 or 12 years and throughout the rest of life).

During the sensorimotor stage, children's thinking is largely realized through their perceptions of the world and their physical interactions with it. Their mental representations are very limited. Consider Piaget's **object permanence task**, which is one of his most famous problems. If an infant younger than 9 months of age is playing with a favorite toy, and another person removes the toy from view, for example by putting it under an opaque cover and not letting the infant immediately reach for it, the infant is very likely to make no effort to retrieve it and to show no emotional distress (Piaget, 1954). This is not due to their being uninterested in the toy or unable to reach for it; if the same toy is put under a clear cover, infants below 9 months readily retrieve it (Munakata, McClelland, Johnson, & Siegler, 1997). Instead, Piaget claimed that infants less than 9 months do not understand that objects continue to exist even when out of sight.

During the preoperational stage, according to Piaget, children can solve not only this simple problem (which they actually can solve after 9 months) but show a wide variety of other symbolic-representation capabilities, such as those involved in drawing and using language. However, such 2- to 7-year-olds tend to focus on a single dimension, even when solving problems would require them to consider multiple dimensions. This is evident in Piaget's (1952) conservation problems. For example, if a glass of water is poured into a taller, thinner glass, children below age 7 generally say that there now is more water than before. Similarly, if a clay ball is reshaped into a long, thin sausage, they claim that there is now more clay, and if a row of coins is spread out, they claim that there are now more coins. In all cases, the children are focusing on one dimension, while ignoring the changes in other dimensions (for example, the greater width of the glass and the clay ball).

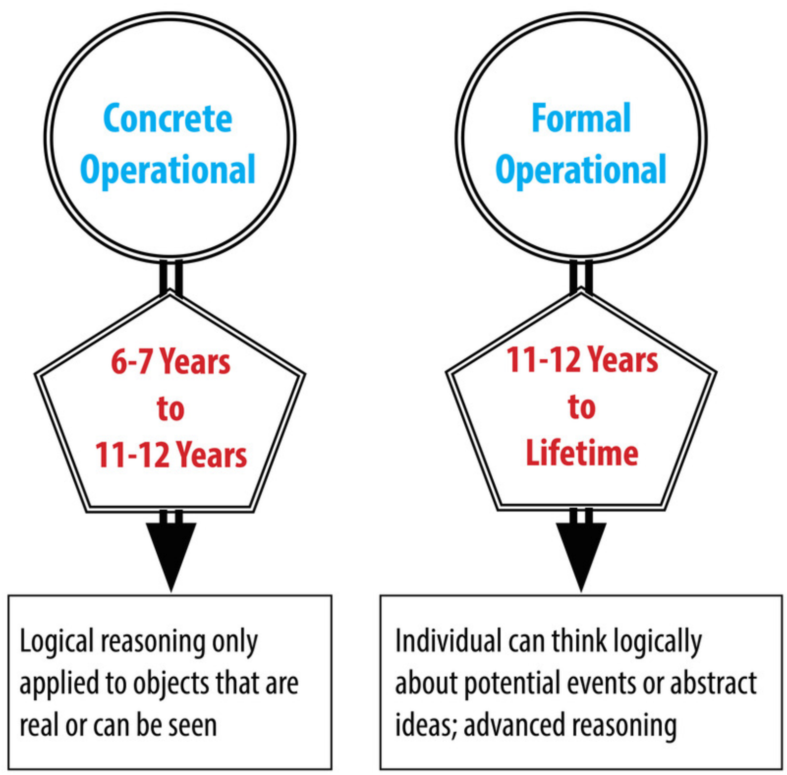


Piaget's Sensorimotor and Pre-operational Reasoning stages

Children overcome this tendency to focus on a single dimension during the concrete operations stage, and think logically in most situations. However, according to Piaget, they still cannot think in systematic scientific ways, even when such thinking would be useful. Thus, if asked to find out which variables influence the period that a pendulum takes to complete its arc, and given weights that they can attach to strings in order to do experiments with the pendulum to find out, most children younger than age 12, perform biased experiments from which no conclusion can be drawn, and then conclude that whatever they originally believed is correct. For example, if a boy believed that weight was the only variable that mattered, he might put the heaviest weight on the shortest string and push it the hardest, and then conclude that just as he thought, weight is the only variable that matters (Inhelder & Piaget, 1958).

Finally, in the formal operations period, children attain the reasoning power of mature adults, which allows them to solve the pendulum problem and a wide range of other problems. However, this formal operations stage tends not to occur without exposure to formal education in scientific reasoning, and appears to be largely or completely absent from some societies that do not provide this type of education.

Although Piaget's theory has been very influential, it has not gone unchallenged. Many more recent researchers have obtained findings indicating that cognitive development is



Piaget's Concrete and Formal Operations stages considerably more continuous than Piaget claimed. For example, Diamond (1985) found that on the object permanence task described above, infants show earlier knowledge if the waiting period is shorter. At age 6 months, they retrieve the hidden object if the wait is no longer than 2 seconds; at 7 months, they retrieve it if the wait is no longer than 4 seconds; and so on. Even earlier, at 3 or 4 months, infants show surprise in the form of longer looking times if objects suddenly appear to vanish with no obvious cause (Baillargeon, 1987). Similarly, children's specific experiences can greatly influence when developmental changes occur. Children of pottery makers in Mexican villages, for example, know that reshaping clay does not change the amount of clay at much younger ages than children who do not have similar experiences (Price-Williams, Gordon, & Ramirez, 1969).

So, is cognitive development fundamentally continuous or fundamentally discontinuous? A reasonable answer seems to be, "It depends on how you look at it and how often you look." For example, under relatively facilitative circumstances, infants show early forms of object permanence by 3 or 4 months, and they gradually extend the range of times for which they can remember hidden objects as they grow older. However, on Piaget's original object permanence task, infants do quite quickly change toward the end of their first year from not reaching for hidden toys to reaching for them, even after they've experienced a substantial delay before being allowed to reach. Thus, the debate between those who emphasize discontinuous, stage-like changes in cognitive development and those who emphasize gradual continuous changes remains a lively one.

Applications to Education

Understanding how children think and learn has proven useful for improving education. One example comes from the area of reading. Cognitive developmental research has shown that **phonemic awareness**—that is, awareness of the component sounds within words—is a crucial skill in learning to read. To measure awareness of the component sounds within words, researchers ask children to decide whether two words rhyme, to decide whether the words start with the same sound, to identify the component sounds within words, and to indicate what would be left if a given sound were removed from a word. Kindergartners' performance on these tasks is the strongest predictor of reading achievement in third and fourth grade, even stronger than IQ or social class background (Nation, 2008). Moreover, teaching these skills to randomly chosen 4- and 5-year-olds results in their being better readers years later (National Reading Panel, 2000).

Another educational application of cognitive developmental research involves the area of mathematics. Even before they enter kindergarten, the mathematical knowledge of children

Activities like playing games that involve working with numbers and spatial relationships can give young children a developmental advantage over peers who have less exposure to the same concepts. [Image: Ben Husmann, https://goo.gl/awOXSw, CC BY 2.0, https://goo.gl/9uSnqN]

from low-income backgrounds lags far behind that of children from more affluent backgrounds. Ramani and Siegler (2008) hypothesized that this difference is due to the children in middle- and upper-income families engaging more frequently in numerical activities, for example playing numerical board games such as Chutes and Ladders. Chutes and Ladders is a game with a number in each square; children start at the number one and spin a spinner or throw a dice to determine how far to move their token. Playing this game seemed likely to teach children about numbers, because in it, larger numbers are associated with greater values on a variety of dimensions. In particular, the higher the number that a child's token reaches, the greater the distance the token will have traveled from the starting point, the greater the number of physical movements the child will have made in moving the token from one square to another, the greater the number of number-words the child will have said and heard, and the more time will have passed since the beginning of the game. These spatial, kinesthetic, verbal, and time-based cues provide a broad-based, multisensory foundation for knowledge of numerical magnitudes (the sizes of numbers), a type of knowledge that is closely related to mathematics achievement test scores (Booth & Siegler, 2006).

Playing this numerical board game for roughly 1 hour, distributed over a 2-week period, improved low-income children's knowledge of numerical magnitudes, ability to read printed numbers, and skill at learning novel arithmetic problems. The gains lasted for months after the game-playing experience (Ramani & Siegler, 2008; Siegler & Ramani, 2009). An advantage of this type of educational intervention is that it has minimal if any cost—a parent could just draw a game on a piece of paper.

Understanding of cognitive development is advancing on many different fronts. One exciting area is linking changes in brain activity to changes in children's thinking (Nelson et al., 2006). Although many people believe that brain maturation is something that occurs before birth, the brain actually continues to change in large ways for many years thereafter. For example, a part of the brain called the prefrontal cortex, which is located at the front of the brain and is particularly involved with planning and flexible problem solving, continues to develop throughout adolescence (Blakemore & Choudhury, 2006). Such new research domains, as well as enduring issues such as nature and nurture, continuity and discontinuity, and how to apply cognitive development research to education, insure that cognitive development will continue to be an exciting area of research in the coming years.

Conclusion

Research into cognitive development has shown us that minds don't just form according to a uniform blueprint or innate intellect, but through a combination of influencing factors. For instance, if we want our kids to have a strong grasp of language we could concentrate on phonemic awareness early on. If we want them to be good at math and science we could engage them in numerical games and activities early on. Perhaps most importantly, we no longer think of brains as empty vessels waiting to be filled up with knowledge but as adaptable organs that develop all the way through early adulthood.

Outside Resources

Book: Frye, D., Baroody, A., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). Teaching math to young children: A practice guide. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.

Book: Goswami, U. G. (2010). The Blackwell Handbook of Childhood Cognitive Development. New York: John Wiley and Sons.

Book: Kuhn, D., & Siegler, R. S. (Vol. Eds.). (2006). Volume 2: Cognition, perception, and language. In W. Damon & R. M. Lerner (Series Eds.), Handbook of child psychology (6th ed.). Hoboken, NJ: Wiley.

Book: Miller, P. H. (2011). Theories of developmental psychology (5th ed.). New York: Worth.

Book: Siegler, R. S., & Alibali, M. W. (2004). Children's thinking (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Discussion Questions

1. Why are there different theories of cognitive development? Why don't researchers agree on which theory is the right one?
2. Do children's natures differ, or do differences among children only reflect differences in their experiences?
3. Do you see development as more continuous or more discontinuous?
4. Can you think of ways other than those described in the module in which research on cognitive development could be used to improve education?

Vocabulary

Chutes and Ladders
A numerical board game that seems to be useful for building numerical knowledge.

Concrete operations stage
Piagetian stage between ages 7 and 12 when children can think logically about concrete situations but not engage in systematic scientific reasoning.

Conservation problems
Problems pioneered by Piaget in which physical transformation of an object or set of objects changes a perceptually salient dimension but not the quantity that is being asked about.

Continuous development
Ways in which development occurs in a gradual incremental manner, rather than through sudden jumps.

Depth perception
The ability to actively perceive the distance from oneself of objects in the environment.

Discontinuous development
Discontinuous development

Formal operations stage
Piagetian stage starting at age 12 years and continuing for the rest of life, in which adolescents may gain the reasoning powers of educated adults.

Information processing theories
Theories that focus on describing the cognitive processes that underlie thinking at any one age and cognitive growth over time.

Nature
The genes that children bring with them to life and that influence all aspects of their development.

Numerical magnitudes
The sizes of numbers.

Nurture
The environments, starting with the womb, that influence all aspects of children's development.

Object permanence task
The Piagetian task in which infants below about 9 months of age fail to search for an object that is removed from their sight and, if not allowed to search immediately for the object, act as if they do not know that it continues to exist.

Phonemic awareness
Awareness of the component sounds within words.

Piaget's theory
Theory that development occurs through a sequence of discontinuous stages: the sensorimotor, preoperational, concrete operational, and formal operational stages.

Preoperational reasoning stage
Period within Piagetian theory from age 2 to 7 years, in which children can represent objects through drawing and language but cannot solve logical reasoning problems, such as the conservation problems.

Qualitative changes
Large, fundamental change, as when a caterpillar changes into a butterfly; stage theories such as Piaget's posit that each stage reflects qualitative change relative to previous stages.

Quantitative changes
Gradual, incremental change, as in the growth of a pine tree's girth.

Sensorimotor stage
Period within Piagetian theory from birth to age 2 years, during which children come to represent the enduring reality of objects.

Sociocultural theories
Theory founded in large part by Lev Vygotsky that emphasizes how other people and the attitudes, values, and beliefs of the surrounding culture influence children's development.

References

Baillargeon, R. (1987). Object permanence in 3 1/2- and 4 1/2-month-old infants. *Developmental Psychology, 23*, 655–664.

Baker, L., Dreher, M. J., & Guthrie, J. T., (Eds.). (2000). *Engaging young readers: Promoting achievement and motivation*. New York: Guilford.

Bartrip, J., Morton, J., & De Schonen, S. (2001). Responses to mother's face in 3-week to 5-month old infants. British *Journal of Developmental Psychology*, 19, 219–232

Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychiatry and Psychology, 47*, 296–312.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*, 189–201.

DeVries, R. (1969). Constancy of genetic identity in the years three to six. Monographs of the Society for Research in Child Development, 34, 127.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development, 56*, 868–883.

Held, R. (1993). What can rates of development tell us about underlying mechanisms? In C. E. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 75–90). Hillsdale, NJ: Erlbaum.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

Langlois, J. H., Ritter, J. M., Casey, R. J., & Sawin, D. B. (1995). Infant attractiveness predicts maternal behaviors and attitudes. *Developmental Psychology*, 31, 464–472.

Munakata, Y., McClelland, J. L., Johnson, M. H., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104*, 686 713.

Nation, K. (2008). Learning to read words. The Quarterly *Journal of Experimental Psychology, 61*, 1121 1133.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute of Child Health and Human Development.

Nelson, C. A., Thomas, K. M., & de Haan, M. (2006). Neural bases of cognitive development. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Volume 2: Cognition, perception, and language* (6th ed., pp. 3–57). Hoboken, NJ: Wiley.

Piaget, J. (1954). *The construction of reality in the child*. New York: BasicBooks.

Piaget, J. (1952). *The child's concept of number*. New York: W. W. Norton.

Price-Williams, D. R., Gordon, W., & Ramirez, M. (1969). Skill and conservation: A study of pottery making children. *Developmental Psychology*, 1, 769.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*, 375–394.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype-environment effects. *Child Development, 54*, 424–435.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*, 545–560.

van den Boom, D. C., & Hoeksma, J. B. (1994). The effect of infant irritability on mother-infant interaction: A growth curve analysis. *Developmental Psychology, 30*, 581–590.

27
Adolescent Development

Jennifer Lansford

Adolescence is a period that begins with puberty and ends with the transition to adulthood (approximately ages 10–20). Physical changes associated with puberty are triggered by hormones. Cognitive changes include improvements in complex and abstract thought, as well as development that happens at different rates in distinct parts of the brain and increases adolescents' propensity for risky behavior because increases in sensation-seeking and reward motivation precede increases in cognitive control. Adolescents' relationships with parents go through a period of redefinition in which adolescents become more autonomous, and aspects of parenting, such as distal monitoring and psychological control, become more salient. Peer relationships are important sources of support and companionship during adolescence yet can also promote problem behaviors. Same-sex peer groups evolve into mixed-sex peer groups, and adolescents' romantic relationships tend to emerge from these groups. Identity formation occurs as adolescents explore and commit to different roles and ideological positions. Nationality, gender, ethnicity, socioeconomic status, religious background, sexual orientation, and genetic factors shape how adolescents behave and how others respond to them, and are sources of diversity in adolescence

Learning Objectives

- Describe major features of physical, cognitive, and social development during adolescence.
- Understand why adolescence is a period of heightened risk taking.
- Be able to explain sources of diversity in adolescent development.

Adolescence Defined

Adolescence is often characterized as a period of transformation, primarily, in terms of physical, cognitive, and social-relational change. [Image: Alex Proimos, https://goo.gl/1jqpnl, CC BY-NC 2.0, https://goo.gl/VnKlK8]

Adolescence is a developmental stage that has been defined as starting with puberty and ending with the transition to adulthood (approximately ages 10–20). Adolescence has evolved historically, with evidence indicating that this stage is lengthening as individuals start puberty earlier and transition to adulthood later than in the past. Puberty today begins, on average, at age 10–11 years for girls and 11–12 years for boys. This average age of onset has decreased gradually over time since the 19th century by 3–4 months per decade, which has been attributed to a range of factors including better nutrition, obesity, increased father absence, and other environmental factors (Steinberg, 2013). Completion of formal education, financial independence from parents, marriage, and parenthood have all been markers of the end of adolescence and beginning of adulthood, and all of these transitions happen, on average, later now than in the past. In fact, the prolonging of adolescence has prompted the introduction of a new developmental period called *emerging adulthood* that captures these developmental changes out of adolescence and into adulthood, occurring from approximately ages 18 to 29 (Arnett, 2000).

This module will outline changes that occur during adolescence in three domains: physical, cognitive, and social. Within the social domain, changes in relationships with parents, peers, and romantic partners will be considered. Next, the module turns to adolescents' psychological and behavioral adjustment, including identity formation, aggression and antisocial behavior, anxiety and depression, and academic achievement. Finally, the module summarizes sources of diversity in adolescents' experiences and development.

Physical Changes

Physical changes of puberty mark the onset of adolescence (Lerner & Steinberg, 2009). For both boys and girls, these changes include a growth spurt in height, growth of pubic and underarm hair, and skin changes (e.g., pimples). Boys also experience growth in facial hair and a deepening of their voice. Girls experience breast development and begin menstruating.

These pubertal changes are driven by hormones, particularly an increase in testosterone for boys and estrogen for girls.

Cognitive Changes

Major changes in the structure and functioning of the brain occur during adolescence and result in cognitive and behavioral developments (Steinberg, 2008). Cognitive changes during adolescence include a shift from concrete to more abstract and complex thinking. Such changes are fostered by improvements during early adolescence in attention, memory, processing speed, and metacognition (ability to think about thinking and therefore make better use of strategies like mnemonic devices that can improve thinking). Early in adolescence, changes in the brain's dopaminergic system contribute to increases in adolescents' sensation-seeking and reward motivation. Later in adolescence, the brain's cognitive control centers in the prefrontal cortex develop, increasing adolescents' self-regulation and future orientation. The difference in timing of the development of these different regions of the brain contributes to more risk taking during middle adolescence because adolescents are motivated to seek thrills that sometimes come from risky behavior, such as reckless driving, smoking, or drinking, and have not yet developed the cognitive control to resist impulses or focus equally on the potential risks (Steinberg, 2008). One of the world's leading experts on adolescent development, Laurence Steinberg, likens this to engaging a powerful engine before the braking system is in place. The result is that adolescents are more prone to risky behaviors than are children or adults.

Dopamine is a neurotransmitter in the brain that produces feelings of pleasure. During adolescence, people tend to do whatever activities produce the most dopamine, without fully considering the consequences of such actions. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Social Changes

Parents

Although peers take on greater importance during adolescence, family relationships remain important too. One of the key changes during adolescence involves a renegotiation of parent–child relationships. As adolescents strive for more independence and autonomy during this time, different aspects of parenting become more salient. For example, parents' distal supervision and monitoring become more important as adolescents spend more time away from parents and in the presence of peers. Parental monitoring encompasses a wide range of behaviors such as parents' attempts to set rules and know their adolescents' friends, activities, and whereabouts, in addition to adolescents' willingness to disclose information to their parents (Stattin & Kerr, 2000). **Psychological control**, which involves manipulation and intrusion into adolescents' emotional and cognitive world through invalidating adolescents' feelings and pressuring them to think in particular ways (Barber, 1996), is another aspect of parenting that becomes more salient during adolescence and is related to more problematic adolescent adjustment.

Peers

Peer relationships are a big part of adolescent development. The influence of peers can be both positive and negative as adolescents experiment together with identity formation and new experiences. [Image: CC0 Public Domain, https://goo.gl/m25gce]

As children become adolescents, they usually begin spending more time with their peers and less time with their families, and these peer interactions are increasingly unsupervised by adults. Children's notions of friendship often focus on shared activities, whereas adolescents' notions of friendship increasingly focus on intimate exchanges of thoughts and feelings. During adolescence, peer groups evolve from primarily single-sex to mixed-sex. Adolescents within a peer group tend to be similar to one another in behavior and attitudes, which has been explained as being a function of **homophily** (adolescents who are similar to one another choose to spend time together in a "birds of a feather flock together" way) and influence (adolescents who spend time together shape each other's behavior and attitudes). One of the most widely studied aspects of adolescent peer influence is known as **deviant peer contagion** (Dishion & Tipsord, 2011), which is the process by which peers reinforce problem behavior by laughing or showing other signs of approval that then increase the likelihood of future problem behavior.

Peers can serve both positive and negative functions during adolescence. Negative peer pressure can lead adolescents to make riskier decisions or engage in more problematic behavior than they would alone or in the presence of their family. For example, adolescents are much more likely to drink alcohol, use drugs, and commit crimes when they are with their friends than when they are alone or with their family. However, peers also serve as an important source of social support and companionship during adolescence, and adolescents with positive peer relationships are happier and better adjusted than those who are socially isolated or have conflictual peer relationships.

Crowds are an emerging level of peer relationships in adolescence. In contrast to friendships (which are reciprocal dyadic relationships) and cliques (which refer to groups of individuals who interact frequently), crowds are characterized more by shared reputations or images than actual interactions (Brown & Larson, 2009). These crowds reflect different prototypic identities (such as jocks or brains) and are often linked with adolescents' social status and peers' perceptions of their values or behaviors.

Romantic relationships

Adolescence is the developmental period during which romantic relationships typically first emerge. Initially, same-sex peer groups that were common during childhood expand into mixed-sex peer groups that are more characteristic of adolescence. Romantic relationships often form in the context of these mixed-sex peer groups (Connolly, Furman, & Konarski, 2000). Although romantic relationships during adolescence are often short-lived rather than long-term committed partnerships, their importance should not be minimized. Adolescents spend a great deal of time focused on romantic relationships, and their positive and negative emotions are more tied to romantic relationships (or lack thereof) than to friendships, family relationships, or school (Furman & Shaffer, 2003). Romantic relationships contribute to adolescents' identity formation, changes in family and peer relationships, and adolescents' emotional and behavioral adjustment.

Furthermore, romantic relationships are centrally connected to adolescents' emerging sexuality. Parents, policymakers, and researchers have devoted a great deal of attention to adolescents' sexuality, in large part because of concerns related to sexual intercourse, contraception, and preventing teen pregnancies. However, sexuality involves more than this narrow focus. For example, adolescence is often when individuals who are lesbian, gay, bisexual, or transgender come to perceive themselves as such (Russell, Clarke, & Clary, 2009).

Thus, romantic relationships are a domain in which adolescents experiment with new behaviors and identities.

Behavioral and Psychological Adjustment

Identity formation

Theories of adolescent development often focus on identity formation as a central issue. For example, in Erikson's (1968) classic theory of developmental stages, identity formation was highlighted as the primary indicator of successful development during adolescence (in contrast to role confusion, which would be an indicator of not successfully meeting the task of adolescence). Marcia (1966) described identify formation during adolescence as involving both decision points and commitments with respect to ideologies (e.g., religion, politics) and occupations. He described four identity statuses: foreclosure, identity diffusion, moratorium, and identity achievement. **Foreclosure** occurs when an individual commits to an identity without exploring options. **Identity diffusion** occurs when adolescents neither explore nor commit to any identities. **Moratorium** is a state in which adolescents are actively exploring options but have not yet made commitments. **Identity achievement** occurs when individuals have explored different options and then made identity commitments. Building on this work, other researchers have investigated more specific aspects of identity. For example, Phinney (1989) proposed a model of ethnic identity development that included stages of unexplored ethnic identity, ethnic identity search, and achieved ethnic identity.

Early, antisocial behavior leads to befriending others who also engage in antisocial behavior, which only perpetuates the downward cycle of aggression and wrongful acts. [Image: Philippe Put, https://goo.gl/14H7HL, CC BY 2.0, https://goo.gl/BRvSA7]

Aggression and antisocial behavior

Several major theories of the development of antisocial behavior treat adolescence as an important period. Patterson's (1982) early versus late starter model of the development of aggressive and antisocial behavior distinguishes youths whose antisocial behavior begins during childhood (early starters) versus adolescence (late starters). According to the theory, early starters are at greater risk for long-term antisocial behavior that extends into adulthood than are late starters. Late starters who become antisocial during adolescence are theorized to experience poor parental monitoring and supervision, aspects of parenting that become more salient during adolescence. Poor monitoring and lack of supervision contribute to increasing involvement with deviant peers, which in turn promotes adolescents' own antisocial behavior. Late starters desist from antisocial behavior when changes in the environment make other options more appealing. Similarly, Moffitt's (1993) life-course persistent versus adolescent-limited model distinguishes between antisocial behavior that begins in childhood versus adolescence. Moffitt regards adolescent-limited antisocial behavior as resulting from a "maturity gap" between adolescents' dependence on and control by adults and their desire to demonstrate their freedom from adult constraint. However, as they continue to develop, and legitimate adult roles and privileges become available to them, there are fewer incentives to engage in antisocial behavior, leading to desistance in these antisocial behaviors.

Anxiety and depression

Developmental models of anxiety and depression also treat adolescence as an important period, especially in terms of the emergence of gender differences in prevalence rates that persist through adulthood (Rudolph, 2009). Starting in early adolescence, compared with males, females have rates of anxiety that are about twice as high and rates of depression that are 1.5 to 3 times as high (American Psychiatric Association, 2013). Although the rates vary across specific anxiety and depression diagnoses, rates for some disorders are markedly higher in adolescence than in childhood or adulthood. For example, prevalence rates for specific phobias are about 5% in children and 3%–5% in adults but 16% in adolescents. Anxiety and depression are particularly concerning because suicide is one of the leading causes of death during adolescence. Developmental models focus on interpersonal contexts in both childhood and adolescence that foster depression and anxiety (e.g., Rudolph, 2009). Family adversity, such as abuse and parental psychopathology, during childhood sets the stage for social and behavioral problems during adolescence. Adolescents with such problems generate stress in their relationships (e.g., by resolving conflict poorly and excessively seeking reassurance) and select into more maladaptive social contexts (e.g., "misery loves company" scenarios in which depressed youths select other depressed youths as friends and then frequently co-ruminate as they discuss their problems, exacerbating negative affect and stress). These processes are intensified for girls compared with boys because girls have more relationship-oriented goals related to intimacy and social approval, leaving them more vulnerable to disruption in these relationships. Anxiety and depression then exacerbate problems in social relationships, which in turn contribute to the stability of anxiety and depression over time.

Academic achievement

Adolescents spend more waking time in school than in any other context (Eccles & Roeser, 2011). Academic achievement during adolescence is predicted by interpersonal (e.g., parental engagement in adolescents' education), intrapersonal (e.g., intrinsic motivation), and institutional (e.g., school quality) factors. Academic achievement is important in its own right as a marker of positive adjustment during adolescence but also because academic achievement sets the stage for future educational and occupational opportunities. The most serious consequence of school failure, particularly dropping out of school, is the high risk of unemployment or underemployment in adulthood that follows. High achievement can set the stage for college or future vocational training and opportunities.

Diversity

Although similar biological changes occur for all adolescents as they enter puberty, these changes can differ significantly depending on one's cultural, ethnic, and societal factors. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Adolescent development does not necessarily follow the same pathway for all individuals. Certain features of adolescence, particularly with respect to biological changes associated with puberty and cognitive changes associated with brain development, are relatively universal. But other features of adolescence depend largely on circumstances that are more environmentally variable. For example, adolescents growing up in one country might have different opportunities for risk taking than adolescents in a different country, and supports and sanctions for different behaviors in adolescence depend on laws and values that might be specific to where adolescents live. Likewise, different cultural norms regarding family and peer relationships shape adolescents' experiences in these domains. For example, in some countries, adolescents' parents are expected to retain control over major decisions, whereas in other countries, adolescents are expected to begin sharing in or taking control of decision making.

Even within the same country, adolescents' gender, ethnicity, immigrant status, religion, sexual orientation, socioeconomic status, and personality can shape both how adolescents behave and how others respond to them, creating diverse developmental contexts for different adolescents. For example, early puberty (that occurs before most other peers have experienced puberty) appears to be associated with worse outcomes for girls than boys, likely in part because girls who enter puberty early tend to associate with older boys, which in turn is associated with early sexual behavior and substance use. For adolescents who are ethnic or sexual minorities, discrimination sometimes presents a set of challenges that nonminorities do not face.

Finally, genetic variations contribute an additional source of diversity in adolescence. Current approaches emphasize gene X environment interactions, which often follow a **differential susceptibility** model (Belsky & Pluess, 2009). That is, particular genetic variations are considered riskier than others, but genetic variations also can make adolescents more or less susceptible to environmental factors. For example, the association between the CHRM2genotype and adolescent externalizing behavior (aggression and delinquency)has been found in adolescents whose parents are low in monitoring behaviors (Dick et al., 2011). Thus, it is important to bear in mind that individual differences play an important role in adolescent development.

Conclusions

Adolescent development is characterized by biological, cognitive, and social changes. Social changes are particularly notable as adolescents become more autonomous from their parents, spend more time with peers, and begin exploring romantic relationships and sexuality. Adjustment during adolescence is reflected in identity formation, which often involves a period of exploration followed by commitments to particular identities. Adolescence is characterized by risky behavior, which is made more likely by changes in the brain in which reward-processing centers develop more rapidly than cognitive control systems, making adolescents more sensitive to rewards than to possible negative consequences. Despite these generalizations, factors such as country of residence, gender, ethnicity, and sexual orientation shape development in ways that lead to diversity of experiences across adolescence.

Outside Resources

Podcasts: Society for Research on Adolescence website with links to podcasts on a variety of topics, from autonomy-relatedness in adolescence, to the health ramifications of growing up in the United States.
http://www.s-r-a.org/sra-news/podcasts

Study: The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a longitudinal study of a nationally representative sample of adolescents in grades 7-12 in the United States during the 1994-95 school year. Add Health combines data on respondents' social, economic, psychological and physical well-being with contextual data on the family, neighborhood, community, school, friendships, peer groups, and romantic relationships.
http://www.cpc.unc.edu/projects/addhealth

Video: This is a series of TED talks on topics from the mysterious workings of the adolescent brain, to videos about surviving anxiety in adolescence.
http://tinyurl.com/lku4a3k

Web: UNICEF website on adolescents around the world. UNICEF provides videos and other resources as part of an initiative to challenge common preconceptions about adolescence.
http://www.unicef.org/adolescence/index.html

Discussion Questions

1. What can parents do to promote their adolescents' positive adjustment?
2. In what ways do changes in brain development and cognition make adolescents particularly susceptible to peer influence?
3. How could interventions designed to prevent or reduce adolescents' problem behavior be developed to take advantage of what we know about adolescent development?
4. Reflecting on your own adolescence, provide examples of times when you think your experience was different from those of your peers as a function of something unique about you.
5. In what ways was your experience of adolescence different from your parents' experience of adolescence? How do you think adolescence may be different 20 years from now?

Vocabulary

Crowds
Adolescent peer groups characterized by shared reputations or images.

Deviant peer contagion
The spread of problem behaviors within groups of adolescents.

Differential susceptibility
Genetic factors that make individuals more or less responsive to environmental experiences.

Foreclosure
Individuals commit to an identity without exploration of options.

Homophily
Adolescents tend to associate with peers who are similar to themselves.

Identity achievement
Individuals have explored different options and then made commitments.

Identity diffusion
Adolescents neither explore nor commit to any roles or ideologies.

Moratorium
State in which adolescents are actively exploring options but have not yet made identity commitments.

Psychological control
Parents' manipulation of and intrusion into adolescents' emotional and cognitive world through invalidating adolescents' feelings and pressuring them to think in particular ways.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*, 469–480.

Barber, B. K. (1996). Parental psychological control: Revisiting a neglected construct. *Child Development, 67*, 3296–3319.

Belsky, J., & Pluess, M. (2009). Beyond diathesis-stress: Differential susceptibility to environmental influences. *Psychological Bulletin, 135*, 885–908.

Brown, B. B., & Larson, J. (2009). Peer relationships in adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (pp. 74–103). New York, NY: Wiley.

Connolly, J., Furman, W., & Konarski, R. (2000). The role of peers in the emergence of heterosexual romantic relationships in adolescence. *Child Development, 71*, 1395–1408.

Dick, D. M., Meyers, J. L., Latendresse, S. J., Creemers, H. E., Lansford, J. E., … Huizink, A. C. (2011). CHRM2, parental monitoring, and adolescent externalizing behavior: Evidence for gene-environment interaction. *Psychological Science, 22*, 481–489.

Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology, 62*, 189–214.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*, 225–241.

Erikson, E. H. (1968). *Identity, youth, and crisis*. New York, NY: Norton.

Furman, W., & Shaffer, L. (2003). The role of romantic relationships in adolescent development. In P. Florsheim (Ed.), *Adolescent romantic relations and sexual behavior: Theory, research, and practical implications* (pp. 3–22). Mahwah, NJ: Erlbaum.

Lerner, R. M., & Steinberg, L. (Eds.). (2009). *Handbook of adolescent psychology*. New York, NY: Wiley.

Marcia, J. E. (1966). Development and validation of ego identity status. *Journal of Personality and Social Psychology, 3*, 551–558.

Moffitt, T. E. (1993). Adolescence-limited and life course persistent antisocial behavior: Developmental taxonomy. *Psychological Review, 100*, 674–701.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Press.

Phinney, J. (1989). Stages of ethnic identity in minority group adolescents. *Journal of Early*

Adolescence, 9, 34–49.

Rudolph, K. D. (2009). The interpersonal context of adolescent depression. In S. Nolen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in adolescents* (pp. 377–418). New York, NY: Taylor and Francis.

Russell, S. T., Clarke, T. J., & Clary, J. (2009). Are teens "post-gay"? Contemporary adolescents' sexual identity labels. *Journal of Youth and Adolescence, 38*, 884–890.

Stattin, H., & Kerr, M. (2000). Parental monitoring: A reinterpretation. *Child Development, 71*, 1072–1085.

Steinberg, L. (2013). *Adolescence* (10th ed.). New York, NY: McGraw-Hill.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*, 78–106.

28
Aging

Tara Queen & Jacqui Smith

Traditionally, research on aging described only the lives of people over age 65 and the very old. Contemporary theories and research recognizes that biogenetic and psychological processes of aging are complex and lifelong. Functioning in each period of life is influenced by what happened earlier and, in turn, affects subsequent change. We all age in specific social and historical contexts. Together, these multiple influences on aging make it difficult to define when middle-age or old age begins. This module describes central concepts and research about adult development and aging. We consider contemporary questions about cognitive aging and changes in personality, self-related beliefs, social relationships, and subjective well-being. These four aspects of psychosocial aging are related to health and longevity

Learning Objectives

- Explain research approaches to studying aging.
- Describe cognitive, psychosocial, and physical changes that occur with age.
- Provide examples of how age-related changes in these domains are observed in the context of everyday life.

Introduction

We are currently living in an aging society (Rowe, 2009). Indeed, by 2030 when the last of the Baby Boomers reach age 65, the U.S. older population will be double that of 2010. Furthermore, because of increases in average life expectancy, each new generation can expect to live longer

Due to positive health trends the population of older adults is increasing steadily. Understanding the psychology of aging will be more important than ever to support this group and help them thrive. [Photo: Woody Hibbard, https://goo.gl/VP7pfz, CC BY 2.0, https://goo.gl/JD0cjj]

than their parents' generation and certainly longer than their grandparents' generation. As a consequence, it is time for individuals of all ages to rethink their personal life plans and consider prospects for a long life. When is the best time to start a family? Will the education gained up to age 20 be sufficient to cope with future technological advances and marketplace needs? What is the right balance between work, family, and leisure throughout life? What's the best age to retire? How can I age successfully and enjoy life to the fullest when I'm 80 or 90? In this module we will discuss several different domains of psychological research on aging that will help answer these important questions.

Overview: Life Span and Life Course Perspectives on Aging

Just as young adults differ from one another, older adults are also not all the same. In each decade of adulthood, we observe substantial heterogeneity in cognitive functioning, personality, social relationships, lifestyle, beliefs, and satisfaction with life. This heterogeneity reflects differences in rates of biogenetic and psychological aging and the sociocultural contexts and history of people's lives (Bronfenbrenner, 1979; Fingerman, Berg, Smith, & Antonucci, 2011). Theories of aging describe how these multiple factors interact and change over time. They describe why functioning differs on average between young, middle-aged, young-old, and very old adults and why there is heterogeneity within these age groups. Life course theories, for example, highlight the effects of social expectations and the normative timing of life events and social roles (e.g., becoming a parent, retirement). They also consider the lifelong cumulative effects of membership in specific cohorts (generations) and sociocultural subgroups (e.g., race, gender, socioeconomic status) and exposure to historical events (e.g., war, revolution, natural disasters; Elder, Johnson, & Crosnoe, 2003; Settersten, 2005). Life span theories complement the life-course perspective with a greater focus on processes within the individual (e.g., the aging brain). This approach emphasizes the patterning of lifelong intra- and inter-individual differences in the shape (gain, maintenance, loss), level, and rate of change (Baltes, 1987, 1997). Both life course and life span researchers generally rely on longitudinal studies to examine hypotheses about different patterns of aging associated with the effects of biogenetic, life history, social, and personal factors. Cross-sectional studies provide information about age-group differences, but these are confounded with cohort, time of study, and historical effects.

Cognitive Aging

Researchers have identified areas of both losses and gains in cognition in older age. Cognitive ability and intelligence are often measured using standardized tests and validated measures. The psychometric approach has identified two categories of intelligence that show different rates of change across the life span (Schaie & Willis, 1996). Fluid intelligence refers to information processing abilities, such as logical reasoning, remembering lists, spatial ability, and reaction time. Crystallized intelligence encompasses abilities that draw upon experience and knowledge. Measures of crystallized intelligence include vocabulary tests, solving number problems, and understanding texts.

There are many stereotypes of older adults. They are sometimes seen as slow because of changes in cognitive processing speed. They are though, on average, excellent at drawing on personal experience and knowledge. And they tend to outperform young adults when it comes to social and emotional challenges. [Image: Alex E. Proimos, https://goo.gl/20SbW8, CC BY-NC 2.0, https://goo.gl/FIlc2e]

With age, systematic declines are observed on cognitive tasks requiring self-initiated, effortful processing, without the aid of supportive memory cues (Park, 2000). Older adults tend to perform poorer than young adults on memory tasks that involve recall of information, where individuals must retrieve information they learned previously without the help of a list of possible choices. For example, older adults may have more difficulty recalling facts such as names or contextual details about where or when something happened (Craik, 2000). What might explain these deficits as we age? As we age, working memory, or our ability to simultaneously store and use information, becomes less efficient (Craik & Bialystok, 2006). The ability to process information quickly also decreases with age. This slowing of processing speed may explain age differences on many different cognitive tasks (Salthouse, 2004). Some researchers have argued that **inhibitory functioning**, or the ability to focus on certain information while suppressing attention to less pertinent information, declines with age and may explain age differences in performance on cognitive tasks (Hasher & Zacks, 1988). Finally, it is well established that our hearing and vision decline as we age. Longitudinal research has proposed that deficits in sensory functioning explain age differences in a variety of cognitive abilities (Baltes & Lindenberger, 1997).

Fewer age differences are observed when memory cues are available, such as for **recognition** memory tasks, or when individuals can draw upon acquired knowledge or experience. For example, older adults often perform as well if not better than young adults on tests of word knowledge or vocabulary. With age often comes expertise, and research has pointed to areas where aging experts perform as well or better than younger individuals. For example, older typists were found to compensate for age-related declines in speed by looking farther ahead at printed text (Salthouse, 1984). Compared to younger players, older chess experts are able to focus on a smaller set of possible moves, leading to greater cognitive efficiency (Charness, 1981). Accrued knowledge of everyday tasks, such as grocery prices, can help older adults to make better decisions than young adults (Tentori, Osheron, Hasher, & May, 2001).

How do changes or maintenance of cognitive ability affect older adults' everyday lives? Researchers have studied cognition in the context of several different everyday activities. One example is driving. Although older adults often have more years of driving experience, cognitive declines related to reaction time or attentional processes may pose limitations under certain circumstances (Park & Gutchess, 2000). Research on interpersonal problem solving suggested that older adults use more effective strategies than younger adults to navigate through social and emotional problems (Blanchard-Fields, 2007). In the context of work, researchers rarely find that older individuals perform poorer on the job (Park & Gutchess, 2000). Similar to everyday problem solving, older workers may develop more efficient strategies and rely on expertise to compensate for cognitive decline.

Personality and Self-Related Processes

Research on adult personality examines normative age-related increases and decreases in the expression of the so-called "Big Five" traits—extraversion, neuroticism, conscientiousness, agreeableness, and openness to new experience. Does personality change throughout adulthood? Previously the answer was no, but contemporary research shows that although some people's personalities are relatively stable over time, others' are not (Lucas & Donnellan, 2011; Roberts & Mroczek, 2008). Longitudinal studies reveal average changes during adulthood in the expression of some traits (e.g., neuroticism and openness decrease with age and conscientiousness increases) and individual differences in these patterns due to idiosyncratic life events (e.g., divorce, illness). Longitudinal research also suggests that adult personality traits, such as conscientiousness, predict important life outcomes including job success, health, and longevity (Friedman, Tucker, Tomlinson-Keasey, Schwartz, Wingard, & Criqui, 1993; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007).

In contrast to the relative stability of personality traits, theories about the aging self-propose changes in self-related knowledge, beliefs, and **autobiographical narratives**. Responses to questions such as “Tell me something about yourself. Who are you?” "What are your hopes for the future?" provide insight into the characteristics and life themes that an individual considers uniquely distinguish him or herself from others. These self-descriptions enhance self-esteem and guide behavior (Markus & Nurius, 1986; McAdams, 2006). Theory suggests that as we age, themes that were relatively unimportant in young and middle adulthood gain in salience (e.g., generativity, health) and that people view themselves as improving over time (Ross & Wilson, 2003). Reorganizing personal life narratives and self-descriptions are the major tasks of midlife and young-old age due to transformations in professional and family roles and obligations. In advanced old age, self-descriptions are often characterized by a life review and reflections about having lived a long life. Birren and Schroots (2006), for example, found the process of life review in late life helped individuals confront and cope with the challenges of old age.

There is a difference between physical age and subjective age as summarized in the saying "You are only as old as you feel." [Image: Emar, CC BY-NC-SA 2.0, https://goo.gl/HEXbAA]

One aspect of the self that particularly interests life span and life course psychologists is the individual’s perception and evaluation of their own aging and identification with an age group. **Subjective age** is a multidimensional construct that indicates how old (or young) a person feels and into which age group a person categorizes him- or herself. After early adulthood, most people say that they feel younger than their chronological age and the gap between subjective age and actual age generally increases. On average, after age 40 people report feeling 20% younger than their actual age (e.g., Rubin & Berntsen, 2006). Asking people how satisfied they are with their own aging assesses an evaluative component of age identity. Whereas some aspects of age identity are positively valued (e.g., acquiring seniority in a profession or becoming a grandparent), others may be less valued, depending on societal context. Perceived physical age (i.e., the age one looks in a mirror) is one aspect that requires considerable self-related adaptation in social and cultural contexts that value young bodies. Feeling younger and being satisfied with one's own aging are expressions of positive self-perceptions of aging. They reflect the operation of self-related processes that enhance well-being. Levy (2009) found that older individuals who are able to adapt to and accept changes in their appearance and physical capacity in a positive way report higher well-being, have better health, and live longer.

Social Relationships

Social ties to family, friends, mentors, and peers are primary resources of information, support, and comfort. Individuals develop and age together with family and friends and interact with others in the community. Across the life course, social ties are accumulated, lost, and transformed. Already in early life, there are multiple sources of heterogeneity in the characteristics of each person's social network of relationships (e.g., size, composition, and quality). Life course and life span theories and research about age-related patterns in social relationships focus on understanding changes in the processes underlying social connections. Antonucci's Convoy Model of Social Relations (2001; Kahn & Antonucci, 1980), for example, suggests that the social connections that people accumulate are held together by exchanges in social support (e.g., tangible and emotional). The frequency, types, and reciprocity of the exchanges change with age and in response to need, and in turn, these exchanges impact the health and well-being of the givers and receivers in the convoy. In many relationships, it is not the actual objective exchange of support that is critical but instead the perception that support is available if needed (Uchino, 2009). Carstensen's Socioemotional Selectivity Theory (1993; Carstensen, Isaacowitz, & Charles, 1999) focuses on changes in motivation for actively seeking social contact with others. She proposes that with increasing age our motivational goals change from information gathering to emotion regulation. To optimize the experience of positive affect, older adults actively restrict their social life to prioritize time spent with emotionally close significant others. In line with this, older marriages are found to be characterized by enhanced positive and reduced negative interactions and older partners show more affectionate behavior during conflict discussions than do middle-aged partners (Carstensen, Gottman, & Levenson, 1995). Research showing that older adults have smaller networks compared to young adults and tend to avoid negative interactions also supports this theory. Similar selective processes are also observed when time horizons for interactions with close partners shrink temporarily for young adults (e.g., impending geographical separations).

Much research focuses on the associations between specific effects of long-term social relationships and health in later life. Older married individuals who receive positive social and emotional support from their partner generally report better health than their unmarried peers (Antonucci, 2001; Umberson, Williams, Powers, Liu, & Needham, 2006; Waite & Gallagher, 2000). Despite the overall positive health effects of being married in old age (compared with being widowed, divorced, or single), living as a couple can have a "dark side" if the relationship is strained or if one partner is the primary caregiver. The consequences of positive and negative aspects of relationships are complex (Birditt & Antonucci, 2008; Rook, 1998; Uchino, 2009). For example, in some circumstances, criticism from a partner may be perceived as valid and useful feedback whereas in others it is considered unwarranted and hurtful. In long-term relationships, habitual negative exchanges might have diminished effects. Parent-child and sibling relationships are often the most long-term and emotion-laden social ties. Across the life span, the parent-child tie, for example, is characterized by a paradox of solidarity, conflict, and ambivalence (Fingerman, Chen, Hay, Cichy, & Lefkowitz, 2006).

Emotion and Well-being

As we get older, the likelihood of losing loved ones or experiencing declines in health increases. Does the experience of such losses result in decreases in well-being in older adulthood? Researchers have found that well-being differs across the life span and that the patterns of these differences depend on how well-being is measured.

Measures of **global subjective well-being** assess individuals' overall perceptions of their lives. This can include questions about life satisfaction or judgments of whether individuals are currently living the best life possible. What factors may contribute to how people respond to these questions? Age, health, personality, social support, and life experiences have been shown to influence judgments of global well-being. It is important to note that predictors of well-being may change as we age. What is important to life satisfaction in young adulthood can be different in later adulthood (George, 2010). Early research on well-being argued that life events such as marriage or divorce can temporarily influence well-being, but people quickly adapt and return to a neutral baseline (called the hedonic treadmill; Diener, Lucas, & Scollon, 2006). More recent research suggests otherwise. Using longitudinal data, researchers have examined well-being prior to, during, and after major life events such as widowhood, marriage, and unemployment (Lucas, 2007). Different life events influence well-being in different ways, and individuals do not often adapt back to baseline levels of well-being. The influence of events, such as unemployment, may have a lasting negative influence on well-being as people age. Research suggests that global well-being is highest in early and later adulthood and lowest in midlife (Stone, Schwartz, Broderick, & Deaton, 2010).

Hedonic well-being refers to the emotional component of well-being and includes measures of positive (e.g., happiness, contentment) and negative affect (e.g., stress, sadness). The pattern of positive affect across the adult life span is similar to that of global well-being, with experiences of positive emotions such as happiness and enjoyment being highest in young and older adulthood. Experiences of negative affect, particularly stress and anger, tend to decrease with age. Experiences of sadness are lowest in early and later adulthood compared to midlife (Stone et al., 2010). Other research finds that older adults report more positive and less negative affect than middle age and younger adults (Magai, 2008; Mroczek, 2001). It should be noted that both global well-being and positive affect tend to taper off during late older adulthood and these declines may be accounted for by increases in health-related losses during these years (Charles & Carstensen, 2010).

Psychological well-being aims to evaluate the positive aspects of psychosocial development, as opposed to factors of ill-being, such as depression or anxiety. Ryff's model of psychological well-being proposes six core dimensions of positive well-being. Older adults tend to report higher environmental mastery (feelings of competence and control in managing everyday life) and autonomy (independence), lower personal growth and purpose in life, and similar levels of positive relations with others as younger individuals (Ryff, 1995). Links between health and interpersonal flourishing, or having high-quality connections with others, may be important in understanding how to optimize quality of life in old age (Ryff & Singer, 2000).

Physical activity is one of the pillars of successful aging. [Image: William Murphy, https://goo.gl/Khsbsb, CC BY-SA 2.0, https://goo.gl/jSSrcO]

Successful Aging and Longevity

Increases in average life expectancy in the 20th century and evidence from twin studies that suggests that genes account for only 25% of the variance in human life spans have opened new questions about implications for individuals and society (Christensen, Doblhammer, Rau, &

Vaupel, 2009). What environmental and behavioral factors contribute to a healthy long life? Is it possible to intervene to slow processes of aging or to minimize cognitive decline, prevent dementia, and ensure life quality at the end of life (Fratiglioni, Paillard-Borg, & Winblad, 2004; Hertzog, Kramer, Wilson, & Lindenberger, 2009; Lang, Baltes, & Wagner, 2007)? Should interventions focus on late life, midlife, or indeed begin in early life? Suggestions that pathological change (e.g., dementia) is not an inevitable component of aging and that pathology could at least be delayed until the very end of life led to theories about **successful aging** and proposals about targets for intervention. Rowe and Kahn (1997) defined three criteria of successful aging: (a) the relative avoidance of disease, disability, and risk factors like high blood pressure, smoking, or obesity; (b) the maintenance of high physical and cognitive functioning; and (c) active engagement in social and productive activities. Although such definitions of successful aging are value-laden, research and behavioral interventions have subsequently been guided by this model. For example, research has suggested that age-related declines in cognitive functioning across the adult life span may be slowed through physical exercise and lifestyle interventions (Kramer & Erickson, 2007). It is recognized, however, that societal and environmental factors also play a role and that there is much room for social change and technical innovation to accommodate the needs of the Baby Boomers and later generations as they age in the next decades.

Outside Resources

Web: Columbia Aging Society
http://www.agingsocietynetwork.org/

Web: Columbia International Longevity Center
http://www.mailman.columbia.edu/academic-departments/centers/columbia-aging/international-longevity-center-knowledge-transfer

Web: National Institute on Aging
http://www.nia.nih.gov/

Web: Stanford Center Longevity
http://longevity3.stanford.edu/

Discussion Questions

1. How do age stereotypes and intergenerational social interactions shape quality of life in older adults? What are the implications of the research of Levy and others?
2. Researchers suggest that there is both stability and change in Big Five personality traits after age 30. What is stable? What changes?
3. Describe the Social Convoy Model of Antonucci. What are the implications of this model for older adults?
4. Memory declines during adulthood. Is this statement correct? What does research show?
5. Is dementia inevitable in old age? What factors are currently thought to be protective?
6. What are the components of successful aging described by Rowe and Kahn (1998) and others? What outcomes are used to evaluate successful aging?

Vocabulary

Age identity
How old or young people feel compared to their chronological age; after early adulthood, most people feel younger than their chronological age.

Autobiographical narratives
A qualitative research method used to understand characteristics and life themes that an individual considers to uniquely distinguish him- or herself from others.

Average life expectancy
Mean number of years that 50% of people in a specific birth cohort are expected to survive. This is typically calculated from birth but is also sometimes re-calculated for people who have already reached a particular age (e.g., 65).

Cohort
Group of people typically born in the same year or historical period, who share common experiences over time; sometimes called a generation (e.g., Baby Boom Generation).

Convoy Model of Social Relations
Theory that proposes that the frequency, types, and reciprocity of social exchanges change with age. These social exchanges impact the health and well-being of the givers and receivers in the convoy.

Cross-sectional studies
Research method that provides information about age group differences; age differences are confounded with cohort differences and effects related to history and time of study.

Crystallized intelligence
Type of intellectual ability that relies on the application of knowledge, experience, and learned information.

Fluid intelligence
Type of intelligence that relies on the ability to use information processing resources to reason logically and solve novel problems.

Global subjective well-being
Individuals’ perceptions of and satisfaction with their lives as a whole.

Hedonic well-being

Component of well-being that refers to emotional experiences, often including measures of positive (e.g., happiness, contentment) and negative affect (e.g., stress, sadness).

Heterogeneity

Inter-individual and subgroup differences in level and rate of change over time.

Inhibitory functioning

Ability to focus on a subset of information while suppressing attention to less relevant information.

Intra- and inter-individual differences

Different patterns of development observed within an individual (intra-) or between individuals (inter-).

Life course theories

Theory of development that highlights the effects of social expectations of age-related life events and social roles; additionally considers the lifelong cumulative effects of membership in specific cohorts and sociocultural subgroups and exposure to historical events.

Life span theories

Theory of development that emphasizes the patterning of lifelong within- and between-person differences in the shape, level, and rate of change trajectories.

Longitudinal studies

Research method that collects information from individuals at multiple time points over time, allowing researchers to track cohort differences in age-related change to determine cumulative effects of different life experiences.

Processing speed

The time it takes individuals to perform cognitive operations (e.g., process information, react to a signal, switch attention from one task to another, find a specific target object in a complex picture).

Psychometric approach

Approach to studying intelligence that examines performance on tests of intellectual functioning.

Recall

Type of memory task where individuals are asked to remember previously learned information without the help of external cues.

Recognition
Type of memory task where individuals are asked to remember previously learned information with the assistance of cues.

Self-perceptions of aging
An individual's perceptions of their own aging process; positive perceptions of aging have been shown to be associated with greater longevity and health.

Social network
Network of people with whom an individual is closely connected; social networks provide emotional, informational, and material support and offer opportunities for social engagement.

Socioemotional Selectivity Theory
Theory proposed to explain the reduction of social partners in older adulthood; posits that older adults focus on meeting emotional over information-gathering goals, and adaptively select social partners who meet this need.

Subjective age
A multidimensional construct that indicates how old (or young) a person feels and into which age group a person categorizes him- or herself

Successful aging
Includes three components: avoiding disease, maintaining high levels of cognitive and physical functioning, and having an actively engaged lifestyle.

Working memory
Memory system that allows for information to be simultaneously stored and utilized or manipulated.

References

Antonucci, T. C. (2001). Social relations: An examination of social networks, social support and sense of control. In J.E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (5th ed., pp. 427–453). New York: Academic Press.

Baltes, P. B . (1987). Theoretical propositions of lifespan developmental psychology: On the dynamics between growth and decline. Developmental Psychology, 23, 611–626.

Baltes, P. B. & Lindenberger, U. (1997). Emergence of powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 12*, 12–21.

Birditt, K., & Antonucci, T. C. (2008). Life sustaining irritations? Relationship quality and mortality in the context of chronic illness. *Social Science & Medicine, 67*, 1291.

Birren, J. E. & Schroots, J. J. F. (2006). Autobiographical memory and the narrative self over the life span. In J. E. Birren and K. Warner Schaie (Eds.) *Handbook of the psychology of aging* (6th ed. pp. 477–499). Burlingham, MA: Elsevier Academic Press.

Blanchard-Fields, F. (2007). Everyday problem solving and emotion: An adult development perspective. Current Directions in Psychological Science, 16, 26–31.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Carstensen, L. L. (1993). Motivation for social contact across the life span: A theory of socioemotional selectivity. In J. E. Jacobs (Ed.), Nebraska Symposium on Motivation, 1992: *Developmental perspectives on motivation* (pp. 209–254). Lincoln, NE: University of Nebraska Press.

Carstensen, L. L., Gottman, J. M., & Levensen, R. W. (1995). Emotional behavior in long-term marriage. *Psychology and Aging, 10*, 140–149.

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist, 54*, 165–181.

Charles, S. T., & Carstensen, L. L. (2010). Social and emotional aging. *Annual review of psychology, 61*, 383–409.

Charness, N. (1981). Search in chess: Age and skill differences.*Journal of Experimental Psychology: Human Perception and Performance, 7*, 467.

Christensen, K., Doblhammer, G., Rau, R., Vaupel, J. W. (2009). Ageing populations: the challenges ahead. *Lancet, 374*, 1196–208.

Craik, F. I., & Bialystok, E. (2006). Cognition through the lifespan: mechanisms of change. *Trends* *in Cognitive Sciences, 10*, 131–138.

Craik, F. I. M. (2000). Age-related changes in human memory. In D. C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 75–92). New York: Psychology Press.

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: revising the adaptation theory of well-being. *American Psychologist, 61*, 305.

Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the Life Course*, 3-19. New York: Kluwer Academic/Plenum Publishers.

Fingerman, K. L., Berg, C. A., Smith, J., & Antonucci, T.C. (Eds.) (2011). *Handbook of Life-Span Development*. New York: Springer

Fingerman, K. L., Chen, P. C., Hay, E., Cichy, K. E., & Lefkowitz, E. S. (2006). Ambivalent reactions in the parent and offspring relationship. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*, 152-160.

Fratiglioni, L., Paillard-Borg, S., & Winblad, B. (2004). An active and socially integrated lifestyle in late life might protect against dementia. Lancet Neurology, 3, 343–353

Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Schwartz, J. E., Wingard, D. L., & Criqui, M. H. (1993). Does childhood personality predict longevity? *Journal of Personality and Social Psychology, 65*, 176–185. doi:10.1037/0022-3514.65.1.176

George, L. K. (2010). Still happy after all these years: Research frontiers on subjective well-being in later life. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 65*, 331.

Hasher, L. & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation*, (Vol. 22, pp. 193–225). San Diego, CA: Academic Press.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2009). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest, 9*, 1–65

Kahn, R. L., & Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles, and social support. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (pp. 253–286). New York: Academic Press.

Kramer, A. F. & Erickson, K. I. (2007). Capitalizing on cortical plasticity: The influence of physical activity on cognition and brain function. *Trends in Cognitive Sciences, 11*, 342–348.

Lang, F. R., Baltes, P. B., & Wagner, G. G. (2007). Desired lifetime and end-of-life desires across ddulthood From 20 to 90: A dual-source information model. *Journal of Gerontology:*

Psychological Sciences, 62B, 268–276.

Levy, B. (2009). Stereotype Embodiment: A Psychosocial Approach to Aging. *Current Directions in Psychological Science, 18(6)*, 332–336

Lucas, R. E. (2007). Adaptation and the Set-Point Model of Subjective Well-Being Does Happiness Change After Major Life Events?. *Current Directions in Psychological Science, 16*, 75–79.

Lucas, R. E. & Donnellan, A. B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology, 101*, 847–861

Magai, C. (2008). Long-lived emotions. In M. Lewis, J. M. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of Emotions*, 376–392. New York: Guilford.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*, 954–969.

McAdams, D. P.(2006). The redemptive self: Generativity and the stories Americans live by. *Research in Human Development, 3(2-3)*, 81–100.

Mroczek, D. K. (2001). Age and emotion in adulthood. *Current Directions in Psychological Science, 10(3)*, 87–90.

Park, D. C. (2000). The basic mechanisms accounting for age-related decline in cognitive function. In D.C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 3–21). New York: Psychology Press.

Park, D. C. & Gutchess, A. H. (2000). Cognitive aging and everyday life. In D.C. Park & N. Schwarz (Eds.), *Cognitive Aging: A Primer* (pp. 217–232). New York: Psychology Press.

Roberts, B. W., & Mroczek, D. K. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*, 31–35.

Roberts, B. W., Kuncel, N., Shiner, R., N., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2(4)*, 313–345. doi:10.1111/j.1745-6916.2007.00047.

Rook, K. S. (1998). Investigating the positive and negative sides of personal relationships: Through a lens darkly? In W. R. Cupach & B. H. Spitzberg (Eds.), *The dark side of close relationships* (pp. 369–393). Mahwah, NJ: Lawrence Erlbaum.

Ross, M., & Wilson, A. E. (2003). Autobiographical memory and conceptions of self: Getting better all the time. *Current Directions in Psychological Science, 12*, 66–69.

Rowe, J. W. (2009). Facts and fictions about an aging America. *Contexts, 8*, 16–21.

Rowe, J. W. & Kahn, R. L. (1997). Successful aging. *The Gerontologist, 37(4)*, 433–440.

Rubin, D., & Berntsen, D. (2006). People over 40 feel 20% younger than their age: Subjective age across the lifespan. *Psychonomic Bulletin & Review. 13*, 776–780

Ryff, C. D. (1995). Psychological well-being in adult life. *Current Directions in Psychological Science, 4(4)*, 99–104

Ryff, C. D., & Singer, B. (2000). Interpersonal flourishing: A positive health agenda for the new millennium. *Personality and Social Psychology Review, 4*, 30–44.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science, 13*, 140–144.

Salthouse, T. A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General, 113*, 345.

Schaie, K. W. & Willis, S. L. (1996). Psychometric intelligence and aging. In F. Blanchard-Fields & T.M. Hess (Eds.), *Perspectives on Cognitive Change in Adulthood and Aging* (pp. 293–322). New York: McGraw Hill.

Settersten, R. A., Jr. (2005). Toward a stronger partnership between lifecourse sociology and life-span psychology. *Research in Human Development, 2(1–2)*, 25–41.

Stone, A. A., Schwartz, J. E., Broderick, J. E., & Deaton, A. (2010). A snapshot of the age distribution of psychological well-being in the United States. *Proceedings of the National Academy of Sciences, 107*, 9985–9990.

Tentori, K., Osherson, D., Hasher, L., & May, C. (2001). Wisdom and aging: Irrational preferences in college students but not older adults. *Cognition, 81*, B87–B96.

Uchino, B. N. (2009). What a lifespan approach might tell us about why distinct measures of social support have differential links to physical health. *Journal of Social and Personal Relationships, 26(1)*, 53–62. DOI: 10.1177/0265407509105521

Umberson, D., Williams, K., Powers, D. A., Liu, H., & Needham, B. (2006). You make me sick: Marital quality and health over the life course. *Journal of Health and Social Behavior, 47*, 1–16.

Waite, L. J. & Gallagher, M. (2000). *The case for marriage: Why married people are happier, healthier, and better off financially*. New York: Doubleday.

Social

29
Social Neuroscience

Tiffany A. Ito & Jennifer T. Kubota

This module provides an overview of the new field of social neuroscience, which combines the use of neuroscience methods and theories to understand how other people influence our thoughts, feelings, and behavior. The module reviews research measuring neural and hormonal responses to understand how we make judgments about other people and react to stress. Through these examples, it illustrates how social neuroscience addresses three different questions: (1) how our understanding of social behavior can be expanded when we consider neural and physiological responses, (2) what the actual biological systems are that implement social behavior (e.g., what specific brain areas are associated with specific social tasks), and (3) how biological systems are impacted by social processes.

Learning Objectives

- Define social neuroscience and describe its three major goals.
- Describe how measures of brain activity such as EEG and fMRI are used to make inferences about social processes.
- Discuss how social categorization occurs.
- Describe how simulation may be used to make inferences about others.
- Discuss the ways in which other people can cause stress and also protect us against stress.

Psychology has a long tradition of using our brains and body to better understand how we think and act. For example, in 1939 Heinrich Kluver and Paul Bucy removed (i.e. lesioned) the temporal lobes in some rhesus monkeys and observed the effect on behavior. Included in these lesions was a subcortical area of the brain called the amygdala. After surgery, the monkeys experienced profound behavioral changes, including loss of fear. These results provided initial evidence that the amygdala plays a role in emotional responses, a finding that has since been confirmed by subsequent studies (Phelps & LeDoux, 2005; Whalen & Phelps, 2009).

What Is Social Neuroscience?

Social neuroscience similarly uses the brain and body to understand how we think and act, with a focus on how we think about and act toward other people. More specifically, we can think of social neuroscience as an interdisciplinary field that uses a range of neuroscience measures to understand how other people influence our thoughts, feelings, and behavior. As such, social neuroscience studies the same topics as social psychology, but does so from a multilevel perspective that includes the study of the brain and body. Figure 1 shows the scope of social neuroscience with respect to the older fields of social psychology and neuroscience. Although the field is relatively new – the term first appeared in 1992 (Cacioppo & Berntson, 1992) – it has grown rapidly, thanks to technological advances making measures

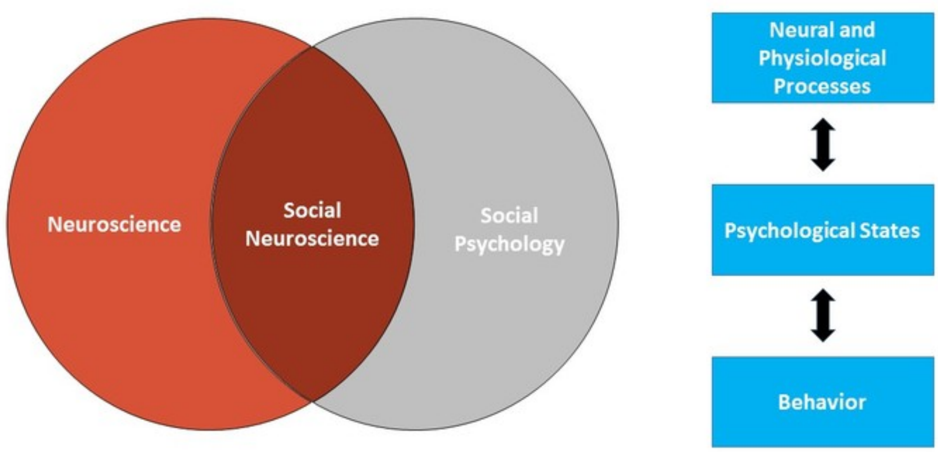


Figure 1. Social neuroscience is the intersection of social psychology and neuroscience. Under this multilevel approach, neural/physiological processes and behavior are two things we can measure or observe. Psychological states cannot be directly observed, but understanding them is the goal. Social neuroscientists use the observable neural/physiological processes and behavioral responses to make inferences about unobservable psychological states. The bidirectional arrows show that all levels of analysis are assumed to influence each other (e.g., psychological states can influence neural responses, and neural responses can influence psychological states).

of the brain and body cheaper and more powerful than ever before, and to the recognition that neural and physiological information are critical to understanding how we interact with other people.

Social neuroscience can be thought of as both a methodological approach (using measures of the brain and body to study social processes) and a theoretical orientation (seeing the benefits of integrating neuroscience into the study of social psychology). The overall approach in social neuroscience is to understand the psychological processes that underlie our social behavior. Because those psychological processes are intrapsychic phenomena that cannot be directly observed, social neuroscientists rely on a combination of measureable or observable neural and physiological responses as well as actual overt behavior to make inferences about psychological states (see Figure 1). Using this approach, social neuroscientists have been able to pursue three different types of questions: (1) What more can we learn about social behavior when we consider neural and physiological responses? (2) What are the actual biological systems that implement social behavior (e.g., what specific brain areas are associated with specific social tasks)? and (3) How are biological systems impacted by social processes?

In this module, we review three research questions that have been addressed with social neuroscience that illustrate the different goals of the field. These examples also expose you to some of the frequently used measures.

How Automatically Do We Judge Other People?

Social categorization is the act of mentally classifying someone as belonging in a group. Why do we do this? It is an effective mental shortcut. Rather than effortfully thinking about every detail of every person we encounter, social categorization allows us to rely on information we already know about the person's group. For example, by classifying your restaurant server as a man, you can quickly activate all the information you have stored about men and use it to guide your behavior. But this shortcut comes with potentially high costs. The stored group beliefs might not be very accurate, and even when they do accurately describe some group members, they are unlikely to be true for every member you encounter. In addition, many beliefs we associate with groups – called **stereotypes** – are negative. This means that relying on social categorization can often lead people to make negative assumptions about others.

The potential costs of social categorization make it important to understand how social categorization occurs. Is it rare or does it occur often? Is it something we can easily stop, or is it hard to override? One difficulty answering these questions is that people are not always consciously aware of what they are doing. In this case, we might not always realize when we are categorizing someone. Another concern is that even when people are aware of their behavior, they can be reluctant to accurately report it to an experimenter. In the case of social categorization, subjects might worry they will look bad if they accurately report classifying someone into a group associated with negative stereotypes. For instance, many racial groups are associated with some negative stereotypes, and subjects may worry that admitting to classifying someone into one of those groups means they believe and use those negative stereotypes.

Figure 2: This man is wearing an elastic electrode cap into which individual electrodes (inside the white circles) are sewn into standardized locations. [Image: Hans, CC0 Public Domain, https://goo.gl/m25gce]

Social neuroscience has been useful for studying how social categorization occurs without having to rely on self-report measures, instead measuring brain activity differences that occur when people encounter members of different social groups. Much of this work has been recorded using the **electroencephalogram**, or **EEG**. EEG is a measure of electrical activity generated by the brain's neurons. Comparing this electrical activity at a given point in time against what a person is thinking and doing at that same time allows us to make inferences about brain activity associated with specific psychological states. One particularly nice feature of EEG is that it provides very precise timing information about when brain activity occurs. EEG is measured non-invasively with small electrodes that rest on the surface of the scalp. This is often done with a stretchy elastic cap, like the one shown in Figure 2, into which the small electrodes are sewn. Researchers simply pull the cap onto the subject's head to get the electrodes into place; wearing it is similar to wearing a swim cap. The subject can then be asked to think about different topics or engage in different tasks as brain activity is measured.

To study social categorization, subjects have been shown pictures of people who belong to different social groups. Brain activity recorded from many individual trials (e.g., looking at lots of different Black individuals) is then averaged together to get an overall idea of how the brain responds when viewing individuals who belong to a particular social group. These studies suggest that social categorization is an **automatic process** – something that happens with little conscious awareness or control – especially for dimensions like gender, race, and age (Ito & Urland, 2003; Mouchetant-Rostaing & Giard, 2003). The studies specifically show that brain activity differs when subjects view members of different social groups (e.g., men versus women, Blacks versus Whites), suggesting that the group differences are being encoded and processed by the perceiver. One interesting finding is that these brain changes occur both when subjects are purposely asked to categorize the people into social groups (e.g., to judge whether the person is Black or White), and also when they are asked to do something that draws attention away from group classifications (e.g., making a personality judgment about the person) (Ito & Urland, 2005). This tells us that we do not have to intend to make group classifications in order for them to happen. It is also very interesting to consider how quickly the changes in brain responses occur. Brain activity is altered by viewing members of different groups within 200 milliseconds of seeing a person's face. That is just two-tenths of a second. Such a fast response lends further support to the idea that social categorization occurs automatically and may not depend on conscious intention.

Overall, this research suggests that we engage in social categorization very frequently. In fact, it appears to happen automatically (i.e., without us consciously intending for it to happen) in most situations for dimensions like gender, age, and race. Since classifying someone into a group is the first step to activating a group stereotype, this research provides important information about how easily stereotypes can be activated. And because it is hard for people to accurately report on things that happen so quickly, this issue has been difficult to study using more traditional self-report measures. Using EEGs has, therefore, been helpful in providing interesting new insights into social behavior.

Do We Use Our Own Behavior to Help Us Understand Others?

Classifying someone into a social group then activating the associated stereotype is one way to make inferences about others. However, it is not the only method. Another strategy is to imagine what our own thoughts, feelings, and behaviors would be in a similar situation. Then we can use our simulated reaction as a best guess about how someone else will respond (Goldman, 2005). After all, we are experts in our own feelings, thoughts, and tendencies. It might be hard to know what other people are feeling and thinking, but we can always ask ourselves how we would feel and act if we were in their shoes.

There has been some debate about whether **simulation** is used to get into the minds of others (Carruthers & Smith, 1996; Gallese & Goldman, 1998). Social neuroscience research has addressed this question by looking at the brain areas used when people think about themselves and others. If the same brain areas are active for the two types of judgments, it lends support to the idea that the self may be used to make inferences about others via simulation.

We know that an area in the prefrontal cortex called the medial prefrontal cortex (mPFC) – located in the middle of the frontal lobe – is active when people think about themselves (Kelley, Macrae, Wyland, Caglar, Inati, & Heatherton, 2002). This conclusion comes from studies using functional magnetic resonance imaging, or fMRI. While EEG measures the brain's electrical activity, fMRI measures changes in the oxygenation of blood flowing in the brain. When neurons become more active, blood flow to the area increases to bring more oxygen and glucose to the active cells. fMRI allows us to image these changes in oxygenation by placing people in an fMRI machine or scanner (Figure 3), which consists of large magnets that create strong magnetic fields. The magnets affect the alignment of the oxygen molecules within the blood (i.e., how they are tilted). As the oxygen molecules move in and out of alignment with the magnetic fields, their nuclei produce energy that can be detected with special sensors

Figure 3. Functional magnetic resonance imaging (fMRI) scanner used to image the brain while people perform tasks. The scanner allows researchers to view the changes in blood oxygenation in specific locations in the brain during a task. Images are collected using powerful magnets and radio waves that shift the position of atoms in oxygenated blood that rushes to areas involved in performing the task. MRIs are non-invasive and there are no known risks from exposure to the magnetic fields or radio waves. Subjects lay on the while bed with their head inside the head coil, then the bed is moved into the scanner. [Image: Janne Moren, https://goo.gl/MKb2jn, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

placed close to the head. Recording fMRI involves having the subject lay on a small bed that is then rolled into the scanner. While fMRI does require subjects to lie still within the small scanner and the large magnets involved are noisy, the scanning itself is safe and painless. Like EEG, the subject can then be asked to think about different topics or engage in different tasks as brain activity is measured. If we know what a person is thinking or doing when fMRI detects a blood flow increase to a particular brain area, we can infer that part of the brain is involved with the thought or action. fMRI is particularly useful for identifying which particular brain areas are active at a given point in time.

The conclusion that the mPFC is associated with the self comes from studies measuring fMRI while subjects think about themselves (e.g., saying whether traits are descriptive of themselves). Using this knowledge, other researchers have looked at whether the same brain area is active when people make inferences about others. Mitchell, Neil Macrae, and Banaji (2005) showed subjects pictures of strangers and had them judge either how pleased the person was to have his or her picture taken or how symmetrical the face appeared. Judging whether someone is pleased about being photographed requires making an inference about someone's internal feelings – we call this **mentalizing**. By contrast, facial symmetry judgments are based solely on physical appearances and do not involve mentalizing. A comparison of brain activity during the two types of judgments shows more activity in the mPFC when making the mental versus physical judgments, suggesting this brain area is involved when inferring the internal beliefs of others.

There are two other notable aspects of this study. First, mentalizing about others also increased activity in a variety of regions important for many aspects of social processing, including a region important in representing biological motion (**superior temporal sulcus** or STS), an area critical for emotional processing (amygdala), and a region also involved in thinking about the beliefs of others (**temporal parietal junction**, TPJ) (Gobbini & Haxby, 2007; Schultz, Imamizu, Kawato, & Frith, 2004) (Figure 4). This finding shows that a distributed and interacting set of brain areas is likely to be involved in social processing. Second, activity in the most ventral part of the mPFC (the part closer to the belly rather than toward the top of the head), which has been most consistently associated with thinking about the self, was particularly active when subjects mentalized about people they rated as similar to themselves. Simulation is thought to be most likely for similar others, so this finding lends support to the conclusion that we use simulation to mentalize about others. After all, if you encounter someone who has the same musical taste as you, you will probably assume you have other things in common with him. By contrast, if you learn that someone loves music that you hate, you might expect him to differ from you in other ways (Srivastava, Guglielmo, & Beer, 2010). Using a simulation of our own feelings and thoughts will be most accurate if we have reason to think the person's internal experiences are like our own. Thus, we may be most likely to use simulation to make inferences about others if we think they are similar to us.

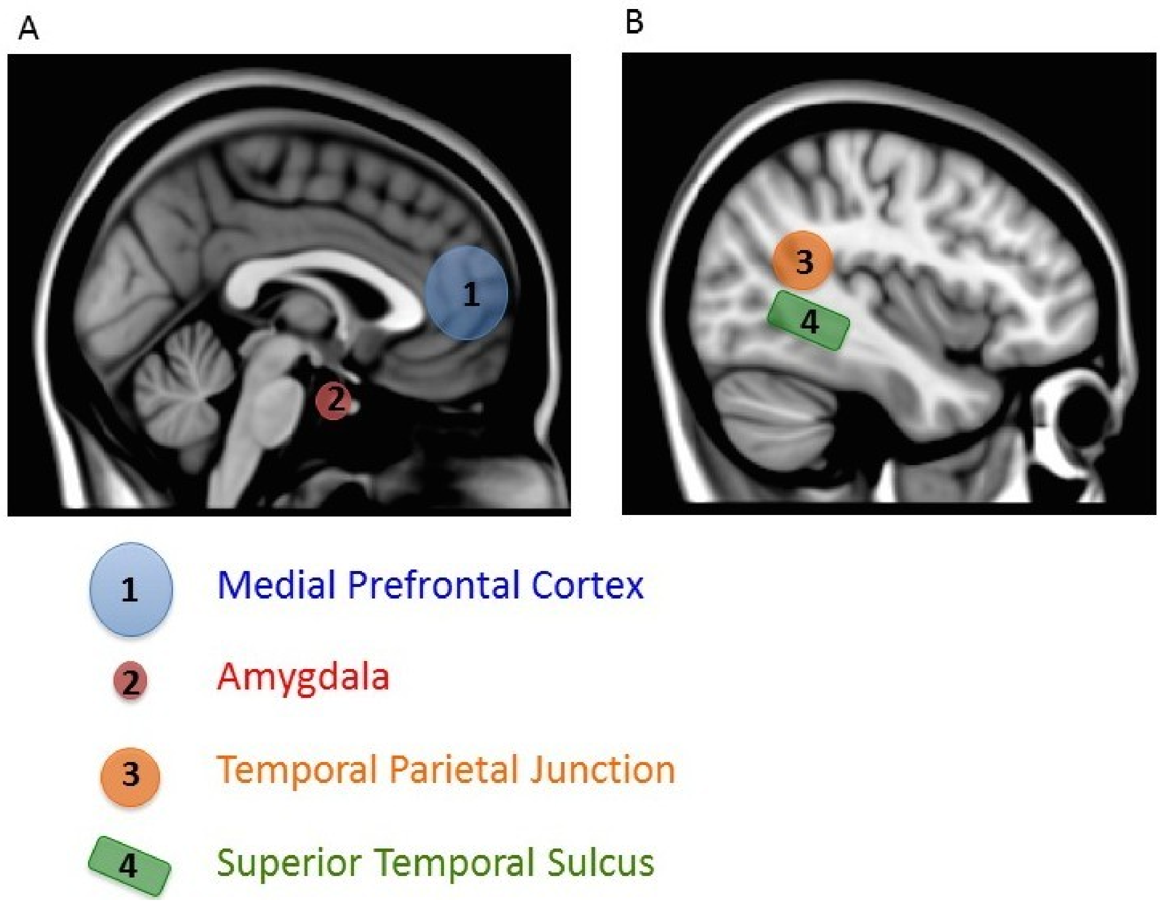


Figure 4. The areas of the brain most commonly associated with processing of self and others. Panel A is a sagittal view, looking at the inside of the brain as if it were sliced in half. Panel B is a lateral view, showing the brain from the outside. The medial prefrontal cortex (mPFC) (Structure 1) is commonly activated when thinking about one's self and when thinking about similar others. Additionally, the amygdala (Structure 2) is important for learning about and detecting important things in our environments and plays an important role in fear learning and expression. The temporal parietal junction (TPJ) (Structure 3), located at the intersection of the parietal and temporal lobes, is activated when people think about the beliefs of others. Activity in the superior temporal sulcus (STS) (Structure 4) is commonly observed when people view biological motion. The common activation of this network of regions when people think about the feelings, thoughts, and intentions of others indicates that the processing of others involves a number of complex psychological processes.

This research is a good example of how social neuroscience is revealing the functional neuroanatomy of social behavior. That is, it tells us which brain areas are involved with social behavior. The mPFC (as well as other areas such as the STS, amygdala, and TPJ) is involved in making judgments about the self and others. This research also provides new information about how inferences are made about others. Whereas some have doubted the widespread use of simulation as a means for making inferences about others, the activation of the mPFC when mentalizing about others, and the sensitivity of this activation to similarity between self and other, provides evidence that simulation occurs.

What Is the Cost of Social Stress?

Stress is an unfortunately frequent experience for many of us. **Stress** – which can be broadly defined as a threat or challenge to our well-being – can result from everyday events like a course exam or more extreme events such as experiencing a natural disaster. When faced with a stressor, **sympathetic nervous system** activity increases in order to prepare our body to respond to the challenge. This produces what Selye (1950) called a **fight or flight response**. The release of **hormones**, which act as messengers from one part of an organism (e.g., a cell or gland) to another part of the organism, is part of the stress response.

A small amount of stress can actually help us stay alert and active. In comparison, sustained stressors, or chronic stress, detrimentally affect our health and impair performance (Al'Absi, Hugdahl, & Lovallo, 2002; Black, 2002; Lazarus, 1974). This happens in part through the chronic secretion of stress-related hormones (e.g., Davidson, Pizzagalli, Nitschke, & Putnam, 2002; Dickerson, Gable, Irwin, Aziz, & Kemeny, 2009). In particular, stress activates the **hypothalamic-pituitary-adrenal (HPA) axis** to release **cortisol** (see Figure 5 for a discussion). Chronic stress, by way of increases in cortisol, impairs attention, memory, and self-control (Arnsten, 2009). Cortisol levels can be measured non-invasively in bodily fluids, including blood and saliva. Researchers often collect a cortisol sample before and after a potentially stressful task. In one common collection method, subjects place polymer swabs under their tongue for 1 to 2 minutes to soak up saliva. The saliva samples are then stored and analyzed later to determine the level of cortisol present at each time point.

Whereas early stress researchers studied the effects of physical stressors like loud noises, social neuroscientists have been instrumental in studying how our interactions with other people can cause stress. This question has been addressed through **neuroendocrinology**, or the study of how the brain and hormones act in concert to coordinate the physiology of the body. One contribution of this work has been in understanding the conditions under which other people can cause stress. In one study, Dickerson, Mycek, and Zaldivar (2008) asked undergraduates to deliver a speech either alone or to two other people. When the students gave the speech in front of others, there was a marked increase in cortisol compared with when they were asked to give a speech alone. This suggests that like chronic physical stress, everyday social stressors, like having your performance judged by others, induces a stress response. Interestingly, simply giving a speech in the same room with someone who is doing something else did not induce a stress response. This suggests that the mere presence of

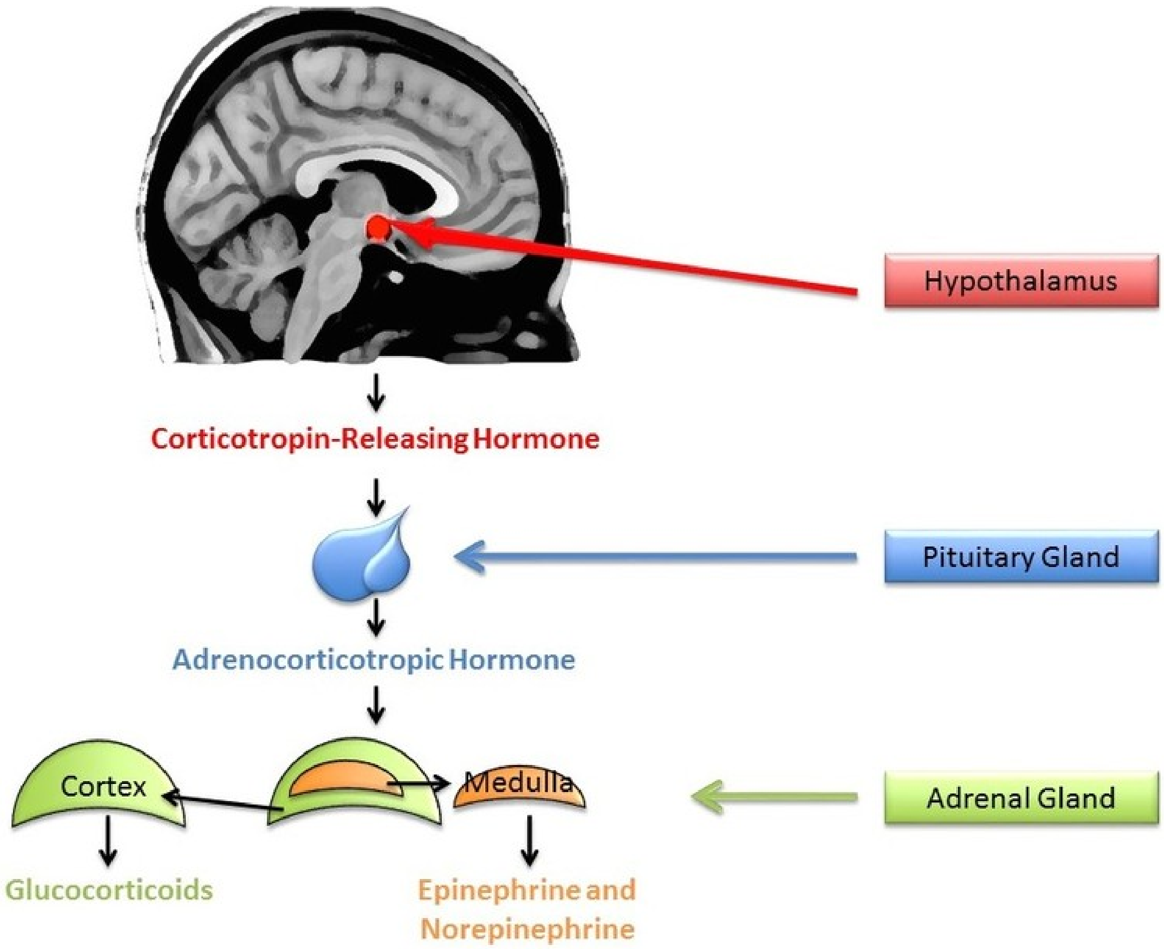


Figure 5: The Hypothalamic-pituitary-adrenal (HPA) axis. Black arrows represent the stress response pathway starting in the brain at the hypothalamus (an area within the brain). Stress triggers neurons in the hypothalamus to release corticotrophin-releasing hormone (CRH). The CRH is transported to the pituitary gland, another area in the brain, that activates the secretion of andrenocorticotropic hormone (ACTH). In turn, ACTH stimulates the adrenal glands that sit on top of the kidneys. The adrenal glands are composed of the outer adrenal cortex and inner adrenal medulla. The adrenal cortex secretes glucorcorticoids (including cortisol) and the medulla secretes epinephrine and norepinephrine. Stress, both psychological and physical, activates the HPA axis and results in the systemic release of cortisol, epinephrine, and norepinephrine.

others is not stressful, but rather it is the potential for them to judge us that induces stress.

Worrying about what other people think of us is not the only source of social stress in our lives. Other research has shown that interacting with people who belong to different social groups than us – what social psychologists call outgroup members – can increase physiological stress responses. For example, cardiovascular responses associated with stress like contractility of the heart ventricles and the amount of blood pumped by the heart (what is called cardiac output) are increased when interacting with outgroup as compared with ingroup members (i.e., people who belong to the same social group we do) (Mendes, Blascovich, Likel, & Hunter, 2002). This stress may derive from the expectation that interactions with dissimilar others will be uncomfortable (Stephan & Stephan, 1985) or concern about being judged as unfriendly and prejudiced if the interaction goes poorly (Plant & Devine, 2003).

The research just reviewed shows that events in our social lives can be stressful, but are social interactions always bad for us? No. In fact, while others can be the source of much stress, they are also a major buffer against stress. Research on **social support** shows that relying on a network of individuals in tough times gives us tools for dealing with stress and can ward off loneliness (Cacioppo & Patrick, 2008). For instance, people who report greater social support show a smaller increase in cortisol when performing a speech in front of two evaluators (Eisenberger, Taylor, Gable, Hilmert, & Lieberman, 2007).

What determines whether others will increase or decrease stress? What matters is the context of the social interaction. When it has potential to reflect badly on the self, social interaction can be stressful, but when it provides support and comfort, social interaction can protect us from the negative effects of stress. Using neuroendocrinology by measuring hormonal changes in the body has helped researchers better understand how social factors impact our body and ultimately our health.

Conclusions

Human beings are intensely social creatures – our lives are intertwined with other people and our health and well-being depend on others. Social neuroscience helps us to understand the critical function of how we make sense of and interact with other people. This module provides an introduction to what social neuroscience is and what we have already learned from it, but there is much still to understand. As we move forward, one exciting future direction will be to better understand how different parts of the brain and body interact to produce the numerous and complex patterns of social behavior that humans display. We hinted at some of this complexity when we reviewed research showing that while the mPFC is involved in mentalizing, other areas such as the STS, amygdala, and TPJ are as well. There are likely additional brain areas involved as well, interacting in ways we do not yet fully understand. These brain areas in turn control other aspects of the body to coordinate our responses during social interactions. Social neuroscience will continue to investigate these questions, revealing new information about how social processes occur, while also increasing our understanding of basic neural and physiological processes.

Outside Resources

Society for Social Neuroscience
http://www.s4sn.org

Video: See a demonstration of fMRI data being collected.
https://www.youtube.com/watch?v=lLORKtkf2n8

Video: See an example of EEG data being collected.
https://www.youtube.com/watch?v=1ovv6lmPHSI

Video: View two tasks frequently used in the lab to create stress – giving a speech in front of strangers, and doing math computations out loud in front of others. Notice how some subjects show obvious signs of stress, but in some situations, cortisol changes suggest that even people who appear calm are experiencing a physiological response associated with stress.
http://www.youtube.com/watch?v=aYI6lCeeT5g

Video: Watch a video used by Fritz Heider and Marianne Simmel in a landmark study on social perception published in 1944. Their goal was to investigate how we perceive other people, and they studied it by seeing how readily we apply people-like interpretations to non-social stimuli.
https://www.youtube.com/watch?v=n9TWwG4SFWQ

Discussion Questions

1. Categorizing someone as a member of a social group can activate group stereotypes. EEG research suggests that social categorization occurs quickly and often automatically. What does this tell us about the likelihood of stereotyping occurring? How can we use this information to develop ways to stop stereotyping from happening?
2. Watch this video, similar to what was used by Fritz Heider and Marianne Simmel in a landmark study on social perception published in 1944, and imagine telling a friend what happened in the video. http://intentionperception.org/wp-content/uploads/2013/02/Heider_Flash.swf. After watching the video, think about the following: Did you describe the motion of the objects solely in geometric terms (e.g., a large triangle moved from the left to the right), or did you describe the movements as actions of animate beings, maybe even of people (e.

g., the circle goes into the house and shuts the door)? In the original research, 33 of 34 subjects described the action of the shapes using human terms. What does this tell us about our tendency to mentalize?

3. Consider the types of things you find stressful. How many of them are social in nature (e.g., are related to your interactions with other people)? Why do you think our social relations have such potential for stress? In what ways can social relations be beneficial and serve as a buffer for stress?

Vocabulary

Amygdala
A region located deep within the brain in the medial area (toward the center) of the temporal lobes (parallel to the ears). If you could draw a line through your eye sloping toward the back of your head and another line between your two ears, the amygdala would be located at the intersection of these lines. The amygdala is involved in detecting relevant stimuli in our environment and has been implicated in emotional responses.

Automatic process
When a thought, feeling, or behavior occurs with little or no mental effort. Typically, automatic processes are described as involuntary or spontaneous, often resulting from a great deal of practice or repetition.

Cortisol
A hormone made by the adrenal glands, within the cortex. Cortisol helps the body maintain blood pressure and immune function. Cortisol increases when the body is under stress.

Electroencephalogram
A measure of electrical activity generated by the brain's neurons.

Fight or flight response
The physiological response that occurs in response to a perceived threat, preparing the body for actions needed to deal with the threat.

Functional magnetic resonance imaging
A measure of changes in the oxygenation of blood flow as areas in the brain become active.

Functional neuroanatomy
Classifying how regions within the nervous system relate to psychology and behavior.

Hormones
Chemicals released by cells in the brain or body that affect cells in other parts of the brain or body.

Hypothalamic-pituitary-adrenal (HPA) axis
A system that involves the hypothalamus (within the brain), the pituitary gland (within the brain), and the adrenal glands (at the top of the kidneys). This system helps maintain

homeostasis (keeping the body's systems within normal ranges) by regulating digestion, immune function, mood, temperature, and energy use. Through this, the HPA regulates the body's response to stress and injury.

Ingroup
A social group to which an individual identifies or belongs.

Lesions
Damage or tissue abnormality due, for example, to an injury, surgery, or a vascular problem.

Medial prefrontal cortex
An area of the brain located in the middle of the frontal lobes (at the front of the head), active when people mentalize about the self and others.

Mentalizing
The act of representing the mental states of oneself and others. Mentalizing allows humans to interpret the intentions, beliefs, and emotional states of others.

Neuroendocrinology
The study of how the brain and hormones act in concert to coordinate the physiology of the body.

Outgroup
A social group to which an individual does not identify or belong.

Simulation
Imaginary or real imitation of other people's behavior or feelings.

Social categorization
The act of mentally classifying someone into a social group (e.g., as female, elderly, a librarian).

Social support
A subjective feeling of psychological or physical comfort provided by family, friends, and others.

Stereotypes
The beliefs or attributes we associate with a specific social group. Stereotyping refers to the act of assuming that because someone is a member of a particular group, he or she possesses the group's attributes. For example, stereotyping occurs when we assume someone is

unemotional just because he is man, or particularly athletic just because she is African American.

Stress
A threat or challenge to our well-being. Stress can have both a psychological component, which consists of our subjective thoughts and feelings about being threatened or challenged, as well as a physiological component, which consists of our body's response to the threat or challenge (see "fight or flight response").

Superior temporal sulcus
The sulcus (a fissure in the surface of the brain) that separates the superior temporal gyrus from the middle temporal gyrus. Located in the temporal lobes (parallel to the ears), it is involved in perception of biological motion or the movement of animate objects.

Sympathetic nervous system
A branch of the autonomic nervous system that controls many of the body's internal organs. Activity of the SNS generally mobilizes the body's fight or flight response.

Temporal parietal junction
The area where the temporal lobes (parallel to the ears) and parieta lobes (at the top of the head toward the back) meet. This area is important in mentalizing and distinguishing between the self and others.

References

Al'Absi, M., Hugdahl, K., & Lovallo, W. (2002). Adrenocortical stress responses and altered working memory performance. *Psychophysiology, 39*(1), 95–99.

Arnsten, A. F. T. (2009). Stress signaling pathways that impair prefrontal cortex structure and function. *Nature Neuroscience Reviews, 10*(6), 410–422.

Black, P. (2002). Stress and the inflammatory response: A review of neurogenic inflammation. * Brain, Behavior, & Immunity, 16*, 622–653.

Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist, 47*, 1019–1028.

Cacioppo, J. T., & Patrick, B. (2008). *Loneliness: Human nature and the need for social connection*. New York, NY: W. W. Norton & Company.

Carruthers, P. and Smith, P. (1996). *Theories of Theories of Mind*. New York, NY: Cambridge University Press.

Davidson, R. J., Pizzagalli, D., Nitschke, J. B., & Putnam, K. (2002). Depression: Perspectives from affective neuroscience. *Annual Review of Psychology, 53*, 545–574.

Dickerson, S. S., Gable, S. L., Irwin, M. R., Aziz, N., & Kemeny, M. E. (2009). Social-evaluative threat and proinflammatory cytokine regulation an experimental laboratory investigation. *Psychological Science, 20*, 1237–1244.

Dickerson, S. S., Mycek, P. J., & Zaldivar, F. (2008). Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychology, 27*(1), 116–121.

Eisenberger, N. I., Taylor, S. E., Gable, S. L., Hilmert, C. J., & Lieberman, M. D. (2007). Neural pathways link social support to attenuated neuroendocrine stress responses. *Neuroimage, 35*(4), 1601–1612.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*, 493–501.

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia, 45*(1), 32–41.

Goldman, A. I. (2005). Imitation, mind reading, and simulation. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science (Vol. 2: Imitation, human development, and culture*, pp. 79–93). Cambridge, MA: MIT Press.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to race and gender of multiply categorizable individuals. *Journal of Personality* *and Social Psychology, 85*, 616–626.

Ito, T.A., & Urland, G.R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, and Behavioral Neuroscience, 5*, 21–36.

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. *Journal of Cognitive Neuroscience, 14*, 785–794.

Lazarus, R. S., (1974). Psychological stress and coping in adaptation and illness. *International Journal of Psychiatry in Medicine, 5*, 321–333.

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interactions with White and Black men. *Personality and Social Psychology Bulletin, 28*, 939–952.

Mitchell, J. P., Neil Macrae, C., & Banaji, M. R. (2005). Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. *Neuroimage, 26*(1), 251–257.

Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *Journal of Cognitive Neuroscience, 15*, 900–910.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron, 48*, 175.

Plant, E. A., & Devine, P. G. (2003). The antecedents and implications of interracial anxiety. * Personality and Social Psychology Bulletin, 29*, 790–801.

Schultz, J., Imamizu, H., Kawato, M., & Frith, C. D. (2004). Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *Journal of Cognitive Neuroscience, 16*, 1695–1705.

Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal: Acta Inc.

Srivastava, S., Guglielmo, S., & Beer, J. S. (2010). Perceiving others' personalities: Examining the dimensionality, assumed similarity to the self, and stability of perceiver effects. *Journal of Personality and Social Psychology, 98*, 520.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues, 41*(3), 157–175.

Whalen, P. J., & Phelps, E. A. (2009). *The human amygdala*. New York, NY: The Guilford Press.

30
Aggression and Violence

Brad J. Bushman

This module discusses the causes and consequences of human aggression and violence. Both internal and external causes are considered. Effective and ineffective techniques for reducing aggression are also discussed.

Learning Objectives

- Explain the important components of the definition of aggression, and explain how aggression differs from violence.
- Explain whether people think the world is less violent now than in the past, and whether it actually is less violent. If there is a discrepancy between perception and reality, how can it be resolved?
- Identify the internal causes and external causes of aggression. Compare and contrast how the inner and external causes differ.
- Identify effective and ineffective approaches to reducing aggression.

Introduction

"*Beware of the dark side. Anger, fear, aggression; the dark side of the Force are they.*"

-Yoda, renowned Jedi master in the *Star Wars* universe

Aggression is indeed the dark side of human nature. Although aggression may have been adaptive in our ancient past, it hardly seems adaptive today. For example, on 14 December 2012 Adam Lanza, age 20, first killed his mother in their home, and then went to an elementary school in Newtown, Connecticut and began shooting, killing 20 children and 6 school employees, before killing himself. When incidents such as these happen, we want to know what caused them. Although it is impossible to know what motivated a particular individual such as Lanza to commit the Newtown school shooting, for decades researchers have studied the internal and external factors that influence aggression and violence. We consider some of these factors in this module.

How much do internal causes such as personality versus external causes such as situations play in aggression? [Image: Dan4th Nicolas, https://goo.gl/RtC4Hi, CC BY 2.0, https://goo.gl/9uSnqN]

Before we get too far, let's begin by defining the term "aggression." Laypeople and researchers often use the term "aggression" differently. Laypeople might describe a salesperson that tries really hard to sell them something as "aggressive." The salesperson does not, however, want to harm potential customers. Most researchers define **aggression** as any behavior intended to harm another person who does not want to be harmed (Baron & Richardson, 1994). This definition includes three important features. First, aggression is a behavior—you can see it. Aggression is not an internal response, such as having angry feelings or aggressive thoughts (although such internal responses can increase the likelihood of actual aggression). Second, aggression is intentional rather than accidental. For example, a dentist might intentionally give a patient a shot of Novocain (which hurts!), but the goal is to help rather than harm the patient. Third, the victim wants to avoid the harm. Thus, suicide and sadomasochistic sex play would not be called aggression because the victim actively seeks to be harmed.

Researchers and laypeople also differ in their use of the term violence. A meteorologist might call a storm "violent" if it has intense winds, rain, thunder, lightning, or hail. Researchers define **violence** as aggression intended to cause extreme physical harm (e.g., injury, death). Thus, all violent acts are aggressive, but not all aggressive acts are violent. For example, screaming and swearing at another person is aggressive, but not violent.

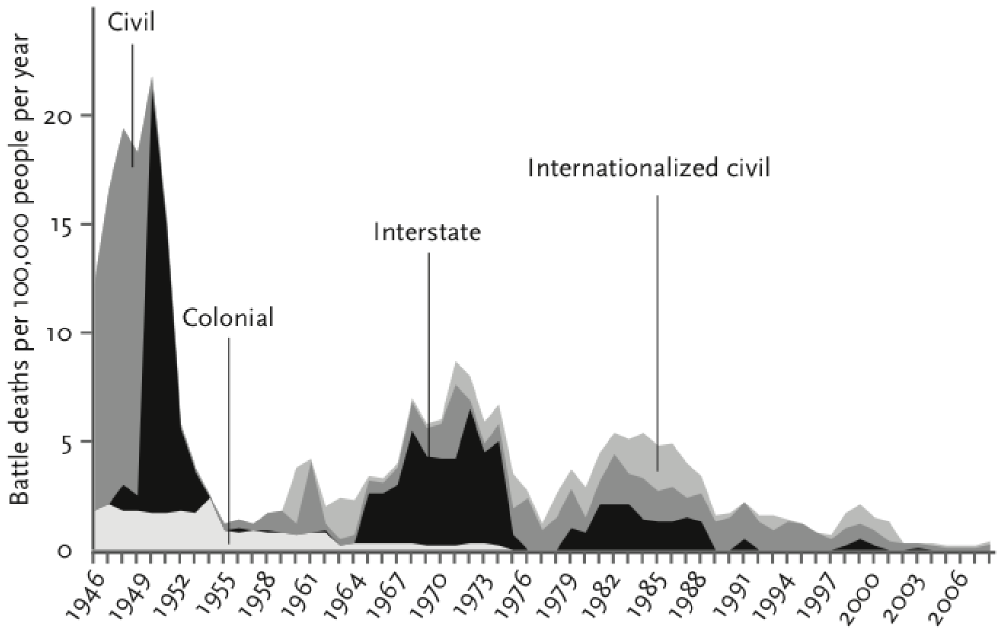


Figure 1. Rate of battle deaths in state-based armed conflicts, 1946-2008. Civilian and military battle deaths in state-based armed conflicts, divided by world population. Sources: UCDP/PRIO Armed Conflict Dataset; see Human Security Report Project (2007), based on data from Lacina and Gleditsch (2005), updated in 2010 by Tara Cooper. "Best" estimate used when available; otherwise the geometric mean of the "High" and "Low" estimates is used. World population figures from U.S. Census Bureau (2010). Population data for 1946-1949 were taken from McEvedy and Jones (1978), and multiplied by 1.01 to make them commensurate with the rest. From Pinker (2011, p. 301). Copyright permission granted by Steven Pinker.

The good news is that the level of violence in the world is decreasing over time—by millennia, century, and even decade (Pinker, 2011). Studies of body counts, such as the proportion of prehistoric skeletons with axe and arrowhead wounds, suggest that prehistoric societies were far more violent than those today. Estimates show that if the wars of the 20th century had killed the same proportion of the population as ancient tribal wars did, then the death toll would have been 20 times higher—2 billion rather than 100 million. More recent data show that murder rates in Europe have decreased dramatically since the Middle Ages. For example, estimated murders in England dropped from 24 per 100,000 in the 14th century to 0.6 per 100,000 by the early 1960s. The major decline in violence occurred in the 17th century during the "Age of Reason," which began in the Netherlands and England and then spread to other European countries. Global violence has also steadily decreased since the middle of the 20th century. For example, the number of battle deaths in interstate wars has declined from more than 65,000 per year in the 1950s to fewer than 2,000 per year in the 2000s. There have also been global declines in the number of armed conflicts and combat deaths, the number of military coups, and the number of deadly violence campaigns waged against civilians. For example, Figure 1 shows the number of battle deaths per 100,000 people per year over 60 years (see Pinker, 2011, p. 301). As can be seen, battle deaths of all types (civil, colonial, interstate, internationalized civil) have decreased over time. The claim that violence has decreased dramatically over time may seem hard to believe in today's digital age when we are constantly bombarded by scenes of violence in the media. In the news media, the top stories are the most violent ones—"If it bleeds it leads," so the saying goes. Citizen journalists around the world also use social media to "show and tell" the world about unjustified acts of violence. Because violent images are more available to us now than ever before, we incorrectly assume that violence levels are also higher. Our tendency to overestimate the amount of violence in the world is due to the **availability heuristic**, which is the tendency to judge the frequency or likelihood of an event by the ease with which relevant instances come to mind. Because we are frequently exposed to scenes of violence in the mass media, acts of violence are readily accessible in memory and come to mind easily, so we assume violence is more common than it actually is.

Human aggression is very complex and is caused by multiple factors. We will consider a few of the most important internal and external causes of aggression. Internal causes include anything the individual brings to the situation that increases the probability of aggression. External causes include anything in the environment that increases the probability of aggression. Finally, we will consider a few strategies for reducing aggression.

Internal Factors

Age

At what age are people most aggressive? You might be surprised to learn that toddlers 1 to 3 years old are most aggressive. Toddlers often rely on physical aggression to resolve conflict and get what they want. In free play situations, researchers have found that 25 percent of their interactions are aggressive (Tremblay, 2000). No other group of individuals (e.g., Mafia, street gangs) resorts to aggression 25 percent of the time. Fortunately for the rest of us, most toddler aggression isn't severe enough to qualify as violence because they don't use weapons, such as guns and knives. As children grow older, they learn to inhibit their aggressive impulses and resolve conflict using nonaggressive means, such as compromise and negotiation. Although most people become less aggressive over time, a small subset of people becomes *more* aggressive over time. The most dangerous years for this small subset of people (and for society as a whole) are late adolescence and early adulthood. For example, 18- to 24-year-olds commit most murders in the U.S. (U.S. Federal Bureau of Investigation, 2012).

Gender

At all ages, males tend to be more physically aggressive than females. However, it would be wrong to think that females are never physically aggressive. Females do use physical aggression, especially when they are provoked by other females (Collins, Quigley, & Leonard, 2007). Among heterosexual partners, women are actually slightly *more* likely than men to use physical aggression (Archer, 2000). However, when men do use physical aggression, they are more likely than women to cause serious injuries and even death to their partners. When people are strongly provoked, gender differences in aggression shrink (Bettencourt & Miller, 1996).

Females are much more likely than males to engage in relational aggression, defined as intentionally harming another person's social relationships, feelings of acceptance, or inclusion within a group (Crick & Grotpeter, 1995). Examples of relational aggression include gossiping, spreading rumors, withdrawing affection to get what you want, excluding someone from your circle of friends, and giving someone the "silent treatment."

Both physical and relational aggression are serious problems in schools and among adolescents. [Image: Elizabet21, https://goo.gl/klf5Pg, CC BY-SA 4.0, https://goo.gl/vUS6LW]

Personality Traits Related to Aggression

Some people seem to be cranky and aggressive almost all the time. Aggressiveness is almost as stable as intelligence over time (Olweus, 1979). Individual differences in aggressiveness are often assessed using self-report questionnaires such as the "Aggression Questionnaire" (Buss & Perry, 1992), which includes items such as "I get into fights a little more than the average person" and "When frustrated, I let my irritation show." Scores on these questionnaires are positively related to actual aggressive and violent behaviors (Anderson & Bushman, 1997).

The components of the "Dark Triad of Personality"—narcissism, psychopathy, and Machiavellianism—are also related to aggression (Paulhus & Williams, 2002). The term "narcissism" comes from the mythical Greek character Narcissus who fell in love with his own image reflected in the water. Narcissists have inflated egos, and they lash out aggressively against others when their inflated egos are threatened (e.g., Bushman & Baumeister, 1998). It is a common myth that aggressive people have low self-esteem (Bushman et al., 2009). Psychopaths are people who lack empathy for others. One of the strongest deterrents of aggression is empathy, which psychopaths lack. The term "Machiavellianism" comes from the Italian philosopher and writer Niccolò Machiavelli, who advocated using any means necessary to gain raw political power, including aggression and violence.

Hostile Cognitive Biases

One key to keeping aggression in check is to give people the benefit of the doubt. Some people, however, do just the opposite. There are three hostile cognitive biases. The **hostile attribution bias** is the tendency to perceive ambiguous actions by others as hostile actions (Dodge, 1980). For example, if a person bumps into you, a hostile attribution would be that the person did it on purpose and wants to hurt you. The **hostile perception bias** is the tendency to perceive social interactions in general as being aggressive (Dill et al., 1997). For example, if you see two people talking in an animated fashion, a hostile perception would be that they are fighting with each other. The **hostile expectation bias** is the tendency to expect others to react to potential conflicts with aggression (Dill et al., 1997). For example, if you bump into another person, a hostile expectation would be that the person will assume that you did it on purpose and will attack you in return. People with hostile cognitive biases view the world as a hostile place.

External Factors

Frustration and Other Unpleasant Events

Are there some situations that are particularly frustrating to you – friends not texting you back, no wi-fi connection available, someone walking at a slow pace in front of you? These situations make make you more likely than usual to behave aggressively. [Image: Syd Daoust, https://goo.gl/Qn9HMu, CC BY-NC-SA 2.0, https://goo.gl/iF4hmM]

One of the earliest theories of aggression proposed that aggression is caused by frustration, which was defined as blocking goal-directed behavior (Dollard et al., 1939). For example, if you are standing in a long line to purchase a ticket, it is frustrating when someone crowds in front of you. This theory was later expanded to say that all unpleasant events, not just frustrations, cause aggression (Berkowitz, 1989). Unpleasant events such as frustrations, provocations, social rejections, hot temperatures, loud noises, bad air (e.g., pollution, foul odors, secondhand smoke), and crowding can all cause aggression. Unpleasant events automatically trigger a fight–flight response.

Alcohol

Alcohol has long been associated with aggression and violence. In fact, sometimes alcohol is deliberately used to promote aggression. It has been standard practice for many centuries to issue soldiers some alcohol before they went into battle, both to increase aggression and reduce fear (Keegan, 1993). There is ample evidence of a link between alcohol and aggression, including evidence from experimental studies showing that consuming alcohol can *cause* an increase in aggression (e.g., Lipsey, Wilson, Cohen, & Derzon, 1997). Most theories of intoxicated aggression fall into one of two categories: (a) pharmacological theories that focus on how alcohol disrupts cognitive processes, and (b) expectancy theories that focus on how social attitudes about alcohol facilitate aggression. Normally, people have strong inhibitions against behaving aggressively, and pharmacological models focus on how alcohol reduces these inhibitions. To use a car analogy, alcohol increases aggression by cutting the brake line rather than by stepping on the gas. How does alcohol cut the brake line? Alcohol disrupts cognitive executive functions that help us organize, plan, achieve goals, and inhibit inappropriate behaviors (Giancola, 2000). Alcohol also reduces glucose, which provides energy to the brain for self-control (Gailliot & Baumeister, 2007). Alcohol has a “myopic” effect on attention—it causes people to focus attention only on the most salient features of a situation and not pay attention to more subtle features (Steele & Josephs, 1990). In some places where alcohol is consumed (e.g., crowded bar), provocations can be salient. Alcohol also reduces self-awareness, which decreases attention to internal standards against behaving aggressively (Hull, 1981).

According to expectancy theories, alcohol increases aggression because people expect it to. In our brains, alcohol and aggression are strongly linked together. Indeed, research shows that subliminally exposing people to alcohol-related words (e.g., vodka) can make them more aggressive, even though they do not drink one drop of alcohol (Subra et al., 2010). In many cultures, drinking occasions are culturally agreed-on “time out” periods where people are not held responsible for their actions (MacAndrew & Edgerton, 1969). Those who behave aggressively when intoxicated sometimes “blame the bottle” for their aggressive actions.

Does this research evidence mean that aggression is somehow contained in alcohol? No. Alcohol increases rather than causes aggressive tendencies. Factors that normally increase aggression (e.g., frustrations and other unpleasant events, aggressive cues) have a stronger effect on intoxicated people than on sober people (Bushman, 1997). In other words, alcohol mainly seems to increase aggression in combination with other factors. If someone insults or attacks you, your response will probably be more aggressive if you are drunk than sober. When there is no provocation, however, the effect of alcohol on aggression may be negligible. Plenty of people enjoy an occasional drink without becoming aggressive.

Reducing Aggression

Most people are greatly concerned about the amount of aggression in society. Aggression directly interferes with our basic needs of safety and security. Thus, it is urgent to find ways to reduce aggression. Because there is no single cause for aggression, it is difficult to design effective treatments. A treatment that works for one individual may not work for another individual. And some extremely aggressive people, such as psychopaths, are considered to be untreatable. Indeed, many people have started to accept the fact that aggression and violence have become an inevitable, intrinsic part of our society. This being said, there certainly are things that can be done to reduce aggression and violence. Before discussing some effective methods for reducing aggression, two ineffective methods need to be debunked: catharsis and punishment.

Catharsis

The term catharsis dates back to Aristotle and means to cleanse or purge. Aristotle taught that viewing tragic plays gave people emotional release from negative emotions. In Greek tragedy, the heroes didn't just grow old and retire—they are often murdered. Sigmund Freud revived the ancient notion of catharsis by proposing that people should express their bottled-up anger. Freud believed if they repressed it, negative emotions would build up inside the individual and surface as psychological disorders. According to catharsis theory, acting aggressively or even viewing aggression purges angry feelings and aggressive impulses into harmless channels. Unfortunately for catharsis theory, research shows the opposite often occurs (e.g., Geen & Quanty, 1977).

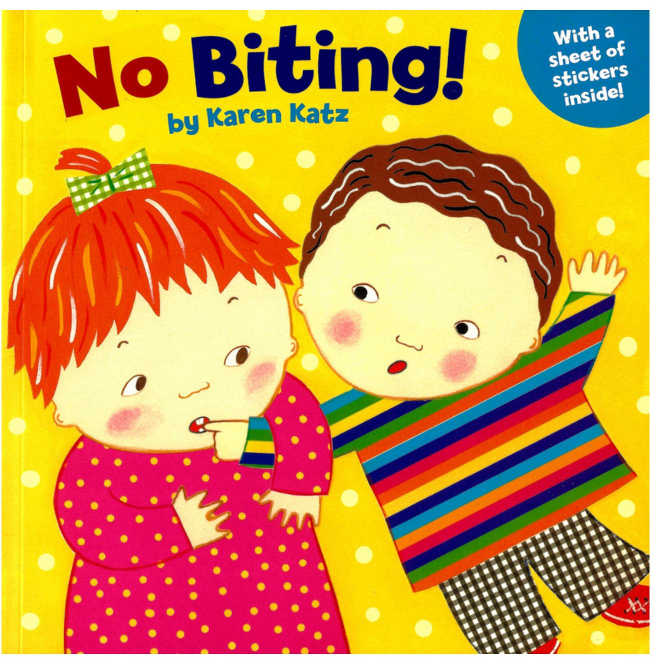
Catharsis is one of the ways to appropriately deal with aggression because it make anger impossible to sustain. [Image: Peter Conlan https://unsplash.com/photos/LEgwEaBVGMo https://unsplash.com/license]

If venting anger doesn't get rid of it, what does? All emotions, including anger, consist of bodily states (e.g., arousal) and mental meanings. To get rid of anger, you can focus on either of those. Anger can be reduced by getting rid of the arousal state, such as by relaxing, listening to calming music, or counting to 10 before responding. Mental tactics can also reduce anger, such as by reframing the situation or by distracting oneself and turning one's attention to more pleasant topics. Incompatible behaviors can also help get rid of anger. For example, petting a puppy, watching a comedy, kissing your lover, or helping someone in need, because those acts are incompatible with anger and, therefore, they make the angry state impossible to sustain (e.g., Baron, 1976). Viewing the provocative situation from a more distant perspective, such as that of a fly on the wall, also helps (Mischkowski, Kross, & Bushman, 2012).

Punishment

Most cultures assume that punishment is an effective way to deter aggression and violence. Punishment is defined as inflicting pain or removing pleasure for a misdeed. Punishment can range in intensity from spanking a child to executing a convicted killer. Parents use it, organizations use it, and governments use it, but does it work? Today, aggression researchers have their doubts. Punishment is most effective when it is: (a) intense, (b) prompt, (c) applied consistently and with certainty, (d) perceived as justified, and (e) possible to replace the undesirable punished behavior with a desirable alternative behavior (Berkowitz, 1993). Even if punishment occurs under these ideal conditions, it may only suppress aggressive behavior temporarily, and it has several undesirable long-term consequences. Most important, punishment models the aggressive behavior it seeks to prevent. Longitudinal studies have shown that children who are physically punished by their parents at home are more aggressive outside the home, such as in school (e.g., Lefkowitz, Huesmann, & Eron, 1978). Because punishment is unpleasant, it can also trigger aggression just like other unpleasant events.

Successful Interventions



One of the ways to circumvent the violent reactions of children who may eventually grow up to be aggressive adults is to model constructive responses to stress and frustration. [Image: Vernon Barford School Library, https://goo.gl/ByOiBc, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Although specific aggression intervention strategies cannot be discussed in any detail here, there are two important general points to be made. First, successful interventions target as many causes of aggression as possible and attempt to tackle them collectively. Interventions that are narrowly focused at removing a single cause of aggression, however well conducted, are bound to fail. In general, external causes are easier to change than internal causes. For example, one can reduce alcohol consumption, and make unpleasant situations more tolerable (e.g., use air conditioners when it is hot, reduce crowding in stressful environments such as prisons and psychiatric wards).

Second, aggression problems are best

treated in early development, when people are still malleable. As was mentioned previously, aggression is very stable over time, almost as stable as intelligence. If young children display excessive levels of aggression (often in the form of hitting, biting, or kicking), it places them at high risk for becoming violent adolescents and even violent adults. It is much more difficult to alter aggressive behaviors when they are part of an adult personality, than when they are still in development.

Yoda warned that anger, fear, and aggression are the dark side of the Force. They are also the dark side of human nature. Fortunately, aggression and violence are decreasing over time, and this trend should continue. We also know a lot more now than ever before about what factors increase aggression and how to treat aggressive behavior problems. When Luke Skywalker was going to enter the dark cave on Degobah (the fictional *Star Wars* planet), Yoda said, "Your weapons, you will not need them." Hopefully, there will come a time in the not-too-distant future when people all over the world will no longer need weapons.

Outside Resources

Book: Bushman, B. J., & Huesmann, L. R. (2010). Aggression. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), Handbook of social psychology (5th ed.) (pp. 833-863). New York: John Wiley & Sons.

TED Talk: Zak Ebrahim
https://www.ted.com/talks/zak_ebrahim_i_am_the_son_of_a_terrorist_here_s_how_i_chose_-peace?language=en#t-528075

Video: From the Inquisitive Mind website, Brad Bushman conducts a short review of terminology and important research concerning aggression and violence.
https://www.youtube.com/watch?v=hGfwflwazJ4

Discussion Questions

1. Discuss whether different examples (hypothetical and real) meet the definition of aggression and the definition of violence.
2. Consider the various causes of aggression described in this module and elsewhere, and discuss whether they can be changed to reduce aggression, and if so how.

Vocabulary

Aggression
Any behavior intended to harm another person who does not want to be harmed.

Availability heuristic
The tendency to judge the frequency or likelihood of an event by the ease with which relevant instances come to mind.

Catharsis
Greek term that means to cleanse or purge. Applied to aggression, catharsis is the belief that acting aggressively or even viewing aggression purges angry feelings and aggressive impulses into harmless channels.

Hostile attribution bias
The tendency to perceive ambiguous actions by others as aggressive.

Hostile expectation bias
The tendency to assume that people will react to potential conflicts with aggression.

Hostile perception bias
The tendency to perceive social interactions in general as being aggressive.

Punishment
Inflicting pain or removing pleasure for a misdeed. Punishment decreases the likelihood that a behavior will be repeated.

Relational aggression
Intentionally harming another person's social relationships, feelings of acceptance, or inclusion within a group.

Violence
Aggression intended to cause extreme physical harm, such as injury or death.

References

Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology, 1* (pp. 19-41).

Archer, J. (2000). Sex differences in aggression between heterosexual partners: A meta-analytic review. *Psychological Bulletin, 126* (pp. 651-680).

Baron, R. A. (1976). The reduction of human aggression: A field study of the influence of incompatible reactions. *Journal of Applied Social Psychology, 6* (pp. 260-274).

Baron, R. A., & Richardson, D. R. (1994). *Human aggression (2nd ed.)*. New York: Plenum Press.

Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin, 106*, 59-73.

Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119* (pp. 422-447).

Bushman, B. J. (1997). Effects of alcohol on human aggression: Validity of proposed explanations. In D. Fuller, R. Dietrich, & E. Gottheil (Eds.), *Recent developments in alcoholism: Alcohol and violence (Vol. 13)* (pp. 227-243). New York: Plenum Press.

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75*, 219-229.

Bushman, B. J., Baumeister, R. F., Thomaes, S., Ryu, E., Begeer, S., & West, S. G. (2009). Looking again, and harder, for a link between low self-esteem and aggression. *Journal of Personality, 77(2)* (pp. 427-446).

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63* (pp. 452-459).

Collins, R. L., Quigley, B., Leonard, K. (2007). Women's physical aggression in bars: An event-based examination of precipitants and predictors of severity. *Aggressive Behavior, 33(4)* (pp. 304-313).

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development, 66* (pp. 710-722).

Dill, K. E., Anderson, C. A., Anderson, K. B., & Deuser, W. E. (1997). Effects of aggressive personality on social expectations and social perceptions. *Journal of Research in Personality, 31* (pp. 272-292).

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development, 51* (pp. 620-635).

Dollard, J., Doob, L., Miller, N., Mowrer, O., & Sears, R. (1939). *Frustration and aggression.* New Haven, CT: Yale University Press.

Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review, 11(4)* (pp. 303-327).

Geen, R. G. & Quanty M. B. (1977). The catharsis of aggression: An evaluation of a hypothesis. In L. Berkowitz (Ed.), *Advances in experimental social psychology (Vol. 10)* (pp. 1-37). New York: Academic Press.

Giancola, P. R. (2000). Executive functioning: A conceptual framework for alcohol-related aggression. *Experimental Clinical Psychopharmacology, 8* (pp. 576-597).

Hull, J. G. (1981). A self-awareness model of the causes and effects of alcohol consumption. *Journal of Abnormal Psychology, 90* (pp. 586-600).

Human Security Report Project (2007). *Human Security Brief 2007.* Vancouver, B.C., Canada: Human Security Report Project

Keegan, J. (1993). *A history of warfare*. New York: Knopf.

Lacina, B., & Gleditsch, N. P. (2005). Monitoring trends in global conflict: A new database in battle deaths. *European Journal of Population, 21*, 145-166.

Lefkowitz, M. M., Huesmann, L. R., & Eron, L. D. (1978). Parental punishment: A longitudinal analysis of effects. *Archives of General Psychiatry, 35(2*), 186-191.

Lipsey, M. W., Wilson, D. B., Cohen M. A., & Derzon, J. H. (1997). Is there a causal relationship between alcohol use and violence? A synthesis of the evidence. In M. Galanter (Ed.), *Recent developments in alcoholism: Vol. 13. Alcohol and violence: Epidemiology, neurobiology, psychology, and family issues,* (pp. 245-282). New York: Plenum Press.

MacAndrew, C., & Edgerton, R. (1969). *Drunken comportment: A social explanation.* Chicago: Aldine.

McEvedy, C., & Jones, R. (1978). *Atlas of world population history.* London: A. Lane.

Mischkowski, D., Kross, E., & Bushman, B. J. (2012). Flies on the wall are less aggressive: Self-distanced reflection reduces angry feelings, aggressive thoughts, and aggressive behaviors. *Journal of Experimental Social Psychology, 48*, 1187-1191.

Olweus, D. (1979). The stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86,* 852-875.

Paulhus, D. L, & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality, 36(6)* (pp. 556-563).

Pinker, S. (2011). *The better angels of our nature*. New York: Viking.

Steele, C. M., & Josephs, R. A. (1990). Alcohol myopia: Its prized and dangerous effects. *American Psychologist, 45* (pp. 921-933).

Subra, B., Muller, D., Bègue, L., Bushman, B. J., & Delmas, F. (2010). Effects of alcohol and weapon cues on aggressive thoughts and behaviors. *Personality and Social Psychology Bulletin, 36(8)* (pp. 1052-1057).

Tremblay, R. E. (2000). The development of aggressive behavior during childhood: What have we learned in the past century? *International Journal of Behavioral Development, 24(2)* (pp. 129-141).

U.S. Census Bureau (2010). International data base (IDB): Total midyear population for the world: 1950-2020. http://www.census.gov/ipc/wwww/idb/worldpop.php

U.S. Federal Bureau of Investigation. (2012). *Uniform crime reports.* Washington, DC: U.S. Government Printing Office.

31
Helping and Prosocial Behavior

Dennis L. Poepsel & David A. Schroeder

People often act to benefit other people, and these acts are examples of prosocial behavior. Such behaviors may come in many guises: helping an individual in need; sharing personal resources; volunteering time, effort, and expertise; cooperating with others to achieve some common goals. The focus of this module is on helping—prosocial acts in dyadic situations in which one person is in need and another provides the necessary assistance to eliminate the other's need. Although people are often in need, help is not always given. Why not? The decision of whether or not to help is not as simple and straightforward as it might seem, and many factors need to be considered by those who might help. In this module, we will try to understand how the decision to help is made by answering the question: Who helps when and why?

Learning Objectives

- Learn which situational and social factors affect when a bystander will help another in need.
- Understand which personality and individual difference factors make some people more likely to help than others.
- Discover whether we help others out of a sense of altruistic concern for the victim, for more self-centered and egoistic motives, or both.

Introduction

Go to YouTube and search for episodes of "Primetime: What Would You Do?" You will find

People often overestimate their willingness to help others in need especially when they are asked about a hypothetical situation rather than encountering one in real life. [Image: Ed Yourdon, https://goo.gl/BYFmcu, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

video segments in which apparently innocent individuals are victimized, while onlookers typically fail to intervene. The events are all staged, but they are very real to the bystanders on the scene. The entertainment offered is the nature of the bystanders' responses, and viewers are outraged when bystanders fail to intervene. They are convinced that they would have helped. But would they? Viewers are overly optimistic in their beliefs that they would play the hero. Helping may occur frequently, but help is not always given to those in need. So *when* do people help, and when do they not? All people are not equally helpful—*who* helps? *Why* would a person help another in the first place? Many factors go into a person's decision to help—a fact that the viewers do not fully appreciate. This module will answer the question: Who helps when and why?

When Do People Help?

Social psychologists are interested in answering this question because it is apparent that people vary in their tendency to help others. In 2010 for instance, Hugo Alfredo Tale-Yax was stabbed when he apparently tried to intervene in an argument between a man and woman. As he lay dying in the street, only one man checked his status, but many others simply glanced at the scene and continued on their way. (One passerby did stop to take a cellphone photo, however.) Unfortunately, failures to come to the aid of someone in need are not unique, as the segments on "What Would You Do?" show. Help is not always forthcoming for those who may need it the most. Trying to understand why people do not always help became the focus of bystander intervention research (e.g., Latané & Darley, 1970).

To answer the question regarding when people help, researchers have focused on

1. how bystanders come to define emergencies,
2. when they decide to take responsibility for helping, and

3. how the costs and benefits of intervening affect their decisions of whether to help.

Defining the situation: The role of pluralistic ignorance

The decision to help is not a simple yes/no proposition. In fact, a series of questions must be addressed before help is given—even in emergencies in which time may be of the essence. Sometimes help comes quickly; an onlooker recently jumped from a Philadelphia subway platform to help a stranger who had fallen on the track. Help was clearly needed and was quickly given. But some situations are ambiguous, and potential helpers may have to decide whether a situation is one in which help, in fact, *needs* to be given.

To define ambiguous situations (including many emergencies), potential helpers may look to the action of others to decide what should be done. But those others are looking around too, also trying to figure out what to do. Everyone is looking, but no one is acting! Relying on others to define the situation and to then erroneously conclude that no intervention is necessary when help is actually needed is called pluralistic ignorance (Latané & Darley, 1970). When people use the *inactions* of others to define their own course of action, the resulting pluralistic ignorance leads to less help being given.

Do I have to be the one to help?: Diffusion of responsibility

How does being in a crowd decrease someone's chance of being helped? How does being in a crowd increase someone's chance of being helped? [Image: flowcomm, https://goo.gl/tiRPch, CC BY 2.0, https://goo.gl/BRvSA7]

Simply being with others may facilitate or inhibit whether we get involved in other ways as well. In situations in which help is needed, the presence or absence of others may affect whether a bystander will assume personal responsibility to give the assistance. If the bystander is alone, personal responsibility to help falls solely on the shoulders of that person. But what if others are present? Although it might seem that having more potential helpers around would increase the chances of the victim getting help, the opposite is often the case. Knowing that someone else *could* help seems to relieve bystanders of personal responsibility, so bystanders do not intervene. This phenomenon is known as **diffusion of responsibility** (Darley & Latané, 1968).

On the other hand, watch the video of the race officials following the 2013 Boston Marathon after two bombs exploded as runners crossed the finish line. Despite the presence of many spectators, the yellow-jacketed race officials immediately rushed to give aid and comfort to the victims of the blast. Each one no doubt felt a personal responsibility to help by virtue of their official capacity in the event; fulfilling the obligations of their roles overrode the influence of the diffusion of responsibility effect.

There is an extensive body of research showing the negative impact of pluralistic ignorance and diffusion of responsibility on helping (Fisher et al., 2011), in both emergencies and everyday need situations. These studies show the tremendous importance potential helpers place on the social situation in which unfortunate events occur, especially when it is not clear what should be done and who should do it. Other people provide important social information about how we should act and what our personal obligations might be. But does knowing a person needs help and accepting responsibility to provide that help mean the person will get assistance? Not necessarily.

The costs and rewards of helping

The nature of the help needed plays a crucial role in determining what happens next. Specifically, potential helpers engage in a **cost–benefit analysis** before getting involved (Dovidio et al., 2006). If the needed help is of relatively low cost in terms of time, money, resources, or risk, then help is more likely to be given. Lending a classmate a pencil is easy; confronting someone who is bullying your friend is an entirely different matter. As the unfortunate case of Hugo Alfredo Tale-Yax demonstrates, intervening may cost the life of the helper.

The potential rewards of helping someone will also enter into the equation, perhaps offsetting the cost of helping. Thanks from the recipient of help may be a sufficient reward. If helpful acts are recognized by others, helpers may receive social rewards of praise or monetary rewards. Even avoiding feelings of guilt if one does not help may be considered a benefit. Potential helpers consider how much helping will cost and compare those costs to the rewards that might be realized; it is the economics of helping. If costs outweigh the rewards, helping is less likely. If rewards are greater than cost, helping is more likely.

Who Helps?

Do you know someone who always seems to be ready, willing, and able to help? Do you know someone who never helps out? It seems there are personality and individual differences in the helpfulness of others. To answer the question of who chooses to help, researchers have examined 1) the role that sex and gender play in helping, 2) what personality traits are associated with helping, and 3) the characteristics of the "prosocial personality."

Who are more helpful—men or women?

In terms of individual differences that might matter, one obvious question is whether men or women are more likely to help. In one of the "What Would You Do?" segments, a man takes a woman's purse from the back of her chair and then leaves the restaurant. Initially, no one responds, but as soon as the woman asks about her missing purse, a group of men immediately rush out the door to catch the thief. So, are men more helpful than women? The quick answer is "not necessarily." It all depends on the type of help needed. To be very clear, the general level of helpfulness may be pretty much equivalent between the sexes, but men and women help in different ways (Becker & Eagly, 2004; Eagly & Crowley, 1986). What accounts for these differences?

Sometimes there are situations that override the gender divide between the helpfulness of men and women and they offer help in equal numbers - for example, volunteering. [Image: Daniel Thornton, https://goo.gl/Rn7yL0, CC BY 2.0, https://goo.gl/BRvSA7]

Two factors help to explain sex and gender differences in helping. The first is related to the cost–benefit analysis process discussed previously. Physical differences between men and women may come into play (e.g., Wood & Eagly, 2002); the fact that men tend to have greater upper body strength than women makes the cost of intervening in some situations less for a man. Confronting a thief is a risky proposition, and some strength may be needed in case the perpetrator decides to fight. A bigger, stronger bystander is less likely to be injured and more likely to be successful.

The second explanation is simple socialization. Men and women have traditionally been raised to play different social roles that prepare them to respond differently to the needs of others, and people tend to help in ways that are most consistent with their gender roles. Female gender roles encourage women to be compassionate, caring, and nurturing; male gender roles encourage men to take physical risks, to be heroic and chivalrous, and to be protective of those less powerful. As a consequence of social training and the gender roles that people have assumed, men may be more likely to jump onto subway tracks to save a fallen passenger, but women are more likely to give comfort to a friend with personal problems (Diekman & Eagly, 2000; Eagly & Crowley, 1986). There may be some specialization in the types of help given by the two sexes, but it is nice to know that there is someone out there—man or woman—who is able to give you the help that you need, regardless of what kind of help it might be.

A trait for being helpful: Agreeableness

Graziano and his colleagues (e.g., Graziano & Tobin, 2009; Graziano, Habishi, Sheese, & Tobin, 2007) have explored how **agreeableness**—one of the Big Five personality dimensions (e.g., Costa & McCrae, 1988)—plays an important role in **prosocial behavior**. Agreeableness is a core trait that includes such dispositional characteristics as being sympathetic, generous, forgiving, and helpful, and behavioral tendencies toward harmonious social relations and likeability. At the conceptual level, a positive relationship between agreeableness and helping may be expected, and research by Graziano et al. (2007) has found that those higher on the agreeableness dimension are, in fact, more likely than those low on agreeableness to help siblings, friends, strangers, or members of some other group. Agreeable people seem to expect that others will be similarly cooperative and generous in interpersonal relations, and they, therefore, act in helpful ways that are likely to elicit positive social interactions.

Searching for the prosocial personality

Rather than focusing on a single trait, Penner and his colleagues (Penner, Fritzsche, Craiger, & Freifeld, 1995; Penner & Orom, 2010) have taken a somewhat broader perspective and identified what they call the **prosocial personality orientation**. Their research indicates that two major characteristics are related to the prosocial personality and prosocial behavior. The first characteristic is called **other-oriented empathy**: People high on this dimension have a strong sense of social responsibility, empathize with and feel emotionally tied to those in need, understand the problems the victim is experiencing, and have a heightened sense of moral obligation to be helpful. This factor has been shown to be highly correlated with the trait of agreeableness discussed previously. The second characteristic, **helpfulness**, is more behaviorally oriented. Those high on the helpfulness factor have been helpful in the past, and because they believe they can be effective with the help they give, they are more likely to be helpful in the future.

Why Help?

Finally, the question of *why* a person would help needs to be asked. What motivation is there for that behavior? Psychologists have suggested that 1) evolutionary forces may serve to predispose humans to help others, 2) egoistic concerns may determine if and when help will be given, and 3) selfless, altruistic motives may also promote helping in some cases.

Evolutionary roots for prosocial behavior

Evolutionary theory suggests that being a good helper was a benefit for survival and reproductive success. And we don't just help our family members, reciprocal altruism has also been a benefit to our survival. [Image: TimJN1, https://goo.gl/iTQfWk, CC BY-SA 2.0, https://goo.gl/eH69he]

Our evolutionary past may provide keys about why we help (Buss, 2004). Our very survival was no doubt promoted by the prosocial relations with clan and family members, and, as a hereditary consequence, we may now be especially likely to help those closest to us—blood-related relatives with whom we share a genetic heritage. According to evolutionary psychology, we are helpful in ways that increase the chances that our DNA will be passed along to future generations (Burnstein, Crandall, & Kitayama, 1994)—the goal of the "selfish gene" (Dawkins, 1976). Our personal DNA may not always move on, but we can still be successful in getting some portion of our DNA transmitted if our daughters, sons, nephews, nieces, and cousins survive to produce offspring. The favoritism shown for helping our blood relatives is called **kin selection** (Hamilton, 1964).

But, we do not restrict our relationships just to our own family members. We live in groups that include individuals who are unrelated to us, and we often help them too. Why? **Reciprocal altruism** (Trivers, 1971) provides the answer. Because of reciprocal altruism, we are all better off in the long run if we help one another. If helping someone now increases the chances that you will be helped later, then your overall chances of survival are increased. There is the chance that someone will take advantage of your help and not return your favors. But people seem predisposed to identify those who fail to reciprocate, and punishments including social exclusion may result (Buss, 2004). Cheaters will not enjoy the benefit of help from others, reducing the likelihood of the survival of themselves and their kin.

Evolutionary forces may provide a general inclination for being helpful, but they may not be as good an explanation for why we help in the here and now. What factors serve as proximal influences for decisions to help?

Egoistic motivation for helping

Most people would like to think that they help others because they are concerned about the other person's plight. In truth, the reasons why we help may be more about ourselves than others: Egoistic or selfish motivations may make us help. Implicitly, we may ask, "What's in it *for me*?" There are two major theories that explain what types of reinforcement helpers may be seeking. The **negative state relief model** (e.g., Cialdini, Darby, & Vincent, 1973; Cialdini, Kenrick, & Baumann, 1982) suggests that people sometimes help in order to make themselves feel better. Whenever we are feeling sad, we can use helping someone else as a positive mood boost to feel happier. Through socialization, we have learned that helping can serve as a secondary reinforcement that will relieve negative moods (Cialdini & Kenrick, 1976).

The **arousal: cost–reward model** provides an additional way to understand why people help (e.g., Piliavin, Dovidio, Gaertner, & Clark, 1981). This model focuses on the aversive feelings aroused by seeing another in need. If you have ever heard an injured puppy yelping in pain, you know that feeling, and you know that the best way to relieve that feeling is to help and to comfort the puppy. Similarly, when we see someone who is suffering in some way (e.g., injured, homeless, hungry), we vicariously experience a sympathetic arousal that is unpleasant, and we are motivated to eliminate that aversive state. One way to do that is to help the person in need. By eliminating the victim's pain, we eliminate our own aversive arousal. Helping is an effective way to alleviate our own discomfort.

As an egoistic model, the arousal: cost–reward model explicitly includes the cost/reward considerations that come into play. Potential helpers will find ways to cope with the aversive arousal that will minimize their costs—maybe by means other than direct involvement. For example, the costs of directly confronting a knife-wielding assailant might stop a bystander from getting involved, but the cost of some *indirect* help (e.g., calling the police) may be acceptable. In either case, the victim's need is addressed. Unfortunately, if the costs of helping are too high, bystanders may reinterpret the situation to justify not helping at all. For some, fleeing the situation causing their distress may do the trick (Piliavin et al., 1981).

The egoistically based negative state relief model and the arousal: cost–reward model see the primary motivation for helping as being the helper's own outcome. Recognize that the victim's outcome is of relatively little concern to the helper—benefits to the victim are incidental byproducts of the exchange (Dovidio et al., 2006). The victim may be helped, but the helper's real motivation according to these two explanations is egoistic: Helpers help to the extent that it makes them feel better.

Altruistic help

Altruism is helping with the aim of improving the wellbeing of others. Having a feeling of empathy for others is an important aspect of altruism. [Image: Ed Yourdon, https://goo.gl/MWCLk1, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Although many researchers believe that egoism is the only motivation for helping, others suggest that altruism—helping that has as its ultimate goal the improvement of another's welfare—may also be a motivation for helping under the right circumstances. Batson (2011) has offered the empathy–altruism model to explain altruistically motivated helping for which the helper expects no benefits. According to this model, the key for altruism is empathizing with the victim, that is, putting oneself in the shoes of the victim and imagining how the victim must feel. When taking this perspective and having empathic concern, potential helpers become primarily interested in increasing the well-being of the victim, even if the helper must incur some costs that might otherwise be easily avoided. The empathy–altruism model does not dismiss egoistic motivations; helpers not empathizing with a victim may experience personal distress and have an egoistic motivation, not unlike the feelings and motivations explained by the arousal: cost–reward model. Because egoistically motivated individuals are primarily concerned with their own cost–benefit outcomes, they are less likely to help if they think they can escape the situation with no costs to themselves. In contrast, altruistically motivated helpers are willing to accept the cost of helping to benefit a person with whom they have empathized—this "self-sacrificial" approach to helping is the hallmark of altruism (Batson, 2011).

Although there is still some controversy about whether people can ever act for purely altruistic motives, it is important to recognize that, while helpers may derive some personal rewards by helping another, the help that has been given is also benefitting someone who was in need. The residents who offered food, blankets, and shelter to stranded runners who were unable to get back to their hotel rooms because of the Boston Marathon bombing undoubtedly received positive rewards because of the help they gave, but those stranded runners who were helped got what they needed badly as well. “In fact, it is quite remarkable how the fates of people who have never met can be so intertwined and complementary. Your benefit is mine; and mine is yours” (Dovidio et al., 2006, p. 143).

Conclusion

Helping feels good to the one who helps and the one who is being helped. [Image: International of Red Cross and Red Crescent Societies, https://goo.gl/0DXo8S, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

We started this module by asking the question, “Who helps when and why?” As we have shown, the question of when help will be given is not quite as simple as the viewers of “What Would You Do?” believe. The power of the situation that operates on potential helpers in real time is not fully considered. What might appear to be a split-second decision to help is actually the result of consideration of multiple situational factors (e.g., the helper’s interpretation of the situation, the presence and ability of others to provide the help, the results of a cost–benefit analysis) (Dovidio et al., 2006). We have found that men and women tend to help in different ways—men are more impulsive and physically active, while women are more nurturing and supportive. Personality characteristics such as agreeableness and the prosocial personality orientation also affect people’s likelihood of giving assistance to others. And, why would people help in the first place? In addition to evolutionary forces (e.g., kin selection, reciprocal altruism), there is extensive evidence to show that helping and prosocial acts may be motivated by selfish, egoistic desires; by selfless, altruistic goals; or by some combination of egoistic and altruistic motives. (For a fuller consideration of the field of prosocial behavior, we refer you to Dovidio et al. [2006].)

Outside Resources

Article: Alden, L. E., & Trew, J. L. (2013). If it makes you happy: Engaging in kind acts increases positive affect in socially anxious individuals. Emotion, 13, 64-75. doi:10.1037/a0027761 Review available at:
http://nymag.com/scienceofus/2015/07/one-way-to-get-over-your-social-anxiety-be-nice.html

Book: Batson, C.D. (2009). Altruism in humans. New York, NY: Oxford University Press.

Book: Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., & Penner, L. A. (2006). The social psychology of prosocial behavior. Mahwah, NJ: Erlbaum.

Book: Mikuliner, M., & Shaver, P. R. (2010). Prosocial motives, emotions, and behavior: The better angels of our nature. Washington, DC: American Psychological Association.

Book: Schroeder, D. A. & Graziano, W. G. (forthcoming). The Oxford handbook of prosocial behavior. New York, NY: Oxford University Press.

Institution: Center for Generosity, University of Notre Dame, 936 Flanner Hall, Notre Dame, IN 46556.
http://www.generosityresearch.nd.edu

Institution: The Greater Good Science Center, University of California, Berkeley.
http://www.greatergood.berkeley.edu

News Article: Bystanders Stop Suicide Attempt
http://jfmueller.faculty.noctrl.edu/crow/bystander.pdf

Social Psychology Network (SPN)
http://www.socialpsychology.org/social.htm#prosocial

Video: Episodes (individual) of "Primetime: What Would You Do?"
http://www.YouTube.com

Video: Episodes of "Primetime: What Would You Do?" that often include some commentary from experts in the field may be available at
http://www.abc.com

Video: From The Inquisitive Mind website, a great overview of different aspects of helping and pro-social behavior including - pluralistic ignorance, diffusion of responsibility, the bystander effect, and empathy.
https://www.youtube.com/watch?v=i2aVjU3F_t0

Discussion Questions

1. Pluralistic ignorance suggests that inactions by other observers of an emergency will decrease the likelihood that help will be given. What do you think will happen if even one other observer begins to offer assistance to a victim?
2. In addition to those mentioned in the module, what other costs and rewards might affect a potential helper's decision of whether to help? Receiving help to solve some problem is an obvious benefit for someone in need; are there any costs that a person might have to bear as a result of receiving help from someone?
3. What are the characteristics possessed by your friends who are most helpful? By your friends who are least helpful? What has made your helpful friends and your unhelpful friends so different? What kinds of help have they given to you, and what kind of help have you given to them? Are you a helpful person?
4. Do you think that sex and gender differences in the frequency of helping and the kinds of helping have changed over time? Why? Do you think that we might expect more changes in the future?
5. What do you think is the primary motive for helping behavior: egoism or altruism? Are there any professions in which people are being "pure" altruists, or are some egoistic motivations always playing a role?
6. There are other prosocial behaviors in addition to the kind of helping discussed here. People volunteer to serve many different causes and organizations. People come together to cooperate with one another to achieve goals that no one individual could reach alone. How do you think the factors that affect helping might affect prosocial actions such as volunteering and cooperating? Do you think that there might be other factors that make people more or less likely to volunteer their time and energy or to cooperate in a group?

Vocabulary

Agreeableness
A core personality trait that includes such dispositional characteristics as being sympathetic, generous, forgiving, and helpful, and behavioral tendencies toward harmonious social relations and likeability.

Altruism
A motivation for helping that has the improvement of another's welfare as its ultimate goal, with no expectation of any benefits for the helper.

Arousal: cost–reward model
An egoistic theory proposed by Piliavin et al. (1981) that claims that seeing a person in need leads to the arousal of unpleasant feelings, and observers are motivated to eliminate that aversive state, often by helping the victim. A cost–reward analysis may lead observers to react in ways other than offering direct assistance, including indirect help, reinterpretation of the situation, or fleeing the scene.

Bystander intervention
The phenomenon whereby people intervene to help others in need even if the other is a complete stranger and the intervention puts the helper at risk.

Cost–benefit analysis
A decision-making process that compares the cost of an action or thing against the expected benefit to help determine the best course of action.

Diffusion of responsibility
When deciding whether to help a person in need, knowing that there are others who could also provide assistance relieves bystanders of some measure of personal responsibility, reducing the likelihood that bystanders will intervene.

Egoism
A motivation for helping that has the improvement of the helper's own circumstances as its primary goal.

Empathic concern
According to Batson's empathy–altruism hypothesis, observers who empathize with a person in need (that is, put themselves in the shoes of the victim and imagine how that person feels)

will experience empathic concern and have an altruistic motivation for helping.

Empathy–altruism model
An altruistic theory proposed by Batson (2011) that claims that people who put themselves in the shoes of a victim and imagining how the victim feel will experience empathic concern that evokes an altruistic motivation for helping.

Helpfulness
A component of the prosocial personality orientation; describes individuals who have been helpful in the past and, because they believe they can be effective with the help they give, are more likely to be helpful in the future.

Helping
Prosocial acts that typically involve situations in which one person is in need and another provides the necessary assistance to eliminate the other's need.

Kin selection
According to evolutionary psychology, the favoritism shown for helping our blood relatives, with the goals of increasing the likelihood that some portion of our DNA will be passed on to future generations.

Negative state relief model
An egoistic theory proposed by Cialdini et al. (1982) that claims that people have learned through socialization that helping can serve as a secondary reinforcement that will relieve negative moods such as sadness.

Other-oriented empathy
A component of the prosocial personality orientation; describes individuals who have a strong sense of social responsibility, empathize with and feel emotionally tied to those in need, understand the problems the victim is experiencing, and have a heightened sense of moral obligations to be helpful.

Personal distress
According to Batson's empathy–altruism hypothesis, observers who take a detached view of a person in need will experience feelings of being "worried" and "upset" and will have an egoistic motivation for helping to relieve that distress.

Pluralistic ignorance
Relying on the actions of others to define an ambiguous need situation and to then erroneously

conclude that no help or intervention is necessary.

Prosocial behavior
Social behavior that benefits another person.

Prosocial personality orientation
A measure of individual differences that identifies two sets of personality characteristics (other-oriented empathy, helpfulness) that are highly correlated with prosocial behavior.

Reciprocal altruism
According to evolutionary psychology, a genetic predisposition for people to help those who have previously helped them.

References

Batson, C. D. (2011). *Altruism in humans*. New York, NY: Oxford University Press.

Becker, S. W., & Eagly, A. H. (2004). The heroism of women and men. *American Psychologist, 59*, 163–178.

Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some neo-Darwinian decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology, 67*, 773–789.

Buss, D. M. (2004). *Evolutionary psychology: The new science of the mind*. Boston, MA: Allyn Bacon.

Cialdini, R. B., & Kenrick, D. T. (1976). Altruism as hedonism: A social developmental perspective on the relationship of negative mood state and helping. *Journal of Personality and Social Psychology, 34*, 907–914.

Cialdini, R. B., Darby, B. K. & Vincent, J. E. (1973). Transgression and altruism: A case for hedonism. *Journal of Experimental Social Psychology, 9*, 502–516.

Cialdini, R. B., Kenrick, D. T., & Baumann, D. J. (1982). Effects of mood on prosocial behavior in children and adults. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 339–359). New York, NY: Academic Press.

Costa, P. T., & McCrae, R. R. (1998). Trait theories in personality. In D. F. Barone, M. Hersen, & V. B. Van Hasselt (Eds.), *Advanced Personality* (pp. 103–121). New York, NY: Plenum.

Darley, J. M. & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*, 377–383.

Dawkins, R. (1976). *The selfish gene*. Oxford, U.K.: Oxford University Press.

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic structures: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin, 26*, 1171–1188.

Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., & Penner, L. A. (2006). *The social psychology of prosocial behavior*. Mahwah, NJ: Erlbaum.

Eagly, A. H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Review, 66*, 183–201.

Fisher, P., Krueger, J. I., Greitemeyer, T., Vogrincie, C., Kastenmiller, A., Frey, D., Henne, M., Wicher, M., & Kainbacher, M. (2011). The bystander-effect: A meta-analytic review of bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*, 517–537.

Graziano, W. G., & Tobin, R. (2009). Agreeableness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of Individual Differences in Social Behavior*. New York, NY: Guilford Press.

Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person x situation perspective. *Journal of Personality and Social Psychology, 93*, 583–599.

Hamilton, W. D. (1964). The genetic evolution of social behavior. *Journal of Theoretical Biology, 7*, 1–52.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York, NY: Appleton-Century-Crofts.

Penner, L. A., & Orom, H. (2010). Enduring goodness: A Person X Situation perspective on prosocial behavior. In M. Mikuliner & P.R. Shaver, P.R. (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature* (pp. 55–72). Washington, DC: American Psychological Association.

Penner, L. A., Fritzsche, B. A., Craiger, J. P., & Freifeld, T. R. (1995). Measuring the prosocial personality. In J. Butcher & C.D. Spielberger (Eds.), *Advances in personality assessment* (Vol. 10, pp. 147–163). Hillsdale, NJ: Erlbaum.

Piliavin, J. A., Dovidio, J. F., Gaertner, S. L., & Clark, R. D., III (1981). *Emergency intervention*. New York, NY: Academic Press.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35–57.

Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin, 128*, 699–727.

Personality

32
Personality Traits

Edward Diener & Richard E. Lucas

Personality traits reflect people's characteristic patterns of thoughts, feelings, and behaviors. Personality traits imply consistency and stability—someone who scores high on a specific trait like Extraversion is expected to be sociable in different situations and over time. Thus, trait psychology rests on the idea that people differ from one another in terms of where they stand on a set of basic trait dimensions that persist over time and across situations. The most widely used system of traits is called the Five-Factor Model. This system includes five broad traits that can be remembered with the acronym OCEAN: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Each of the major traits from the Big Five can be divided into facets to give a more fine-grained analysis of someone's personality. In addition, some trait theorists argue that there are other traits that cannot be completely captured by the Five-Factor Model. Critics of the trait concept argue that people do not act consistently from one situation to the next and that people are very influenced by situational forces. Thus, one major debate in the field concerns the relative power of people's traits versus the situations in which they find themselves as predictors of their behavior.

Learning Objectives

- List and describe the "Big Five" ("OCEAN") personality traits that comprise the Five-Factor Model of personality.
- Describe how the facet approach extends broad personality traits.
- Explain a critique of the personality-trait concept.
- Describe in what ways personality traits may be manifested in everyday behavior.
- Describe each of the Big Five personality traits, and the low and high end of the dimension.
- Give examples of each of the Big Five personality traits, including both a low and high example.

- Describe how traits and social learning combine to predict your social activities.
- Describe your theory of how personality traits get refined by social learning.

Introduction

When we observe people around us, one of the first things that strikes us is how different people are from one another. Some people are very talkative while others are very quiet. Some are active whereas others are couch potatoes. Some worry a lot, others almost never seem anxious. Each time we use one of these words, words like "talkative," "quiet," "active," or "anxious," to describe those around us, we are talking about a person's **personality**—the characteristic ways that people differ from one another. Personality psychologists try to describe and understand these differences.

Although there are many ways to think about the personalities that people have, Gordon Allport and other "personologists" claimed that we can best understand the differences between individuals by understanding their personality traits. **Personality traits** reflect basic dimensions on which people differ (Matthews, Deary, & Whiteman, 2003). According to trait psychologists, there are a limited number of these dimensions (dimensions like Extraversion, Conscientiousness, or Agreeableness), and each individual falls somewhere on each dimension, meaning that they could be low, medium, or high on any specific trait.

"Are you an introvert"? In popular culture it's common to talk about people being introverts or extroverts as if these were precise descriptions that meant the same thing for everyone. But research shows that these traits and others are quite variable within individuals. [Image: Nguyen Hung Vu, https://goo.gl/qKJUAC, CC BY 2.0, https://goo.gl/BRvSA7]

An important feature of personality traits is that they reflect **continuous distributions** rather than distinct personality types. This means that when personality psychologists talk about Introverts and Extraverts, they are not really talking about two distinct types of people who are completely and qualitatively different from one another. Instead, they are talking about people who score relatively low or relatively high along a continuous distribution. In fact, when personality psychologists measure traits like Extraversion, they typically find that most people score somewhere in the middle, with smaller numbers showing more extreme levels. The figure below shows the distribution of Extraversion scores from a survey of thousands of people. As you can see, most people report being moderately, but not extremely, extraverted, with fewer people reporting very high or very low scores.

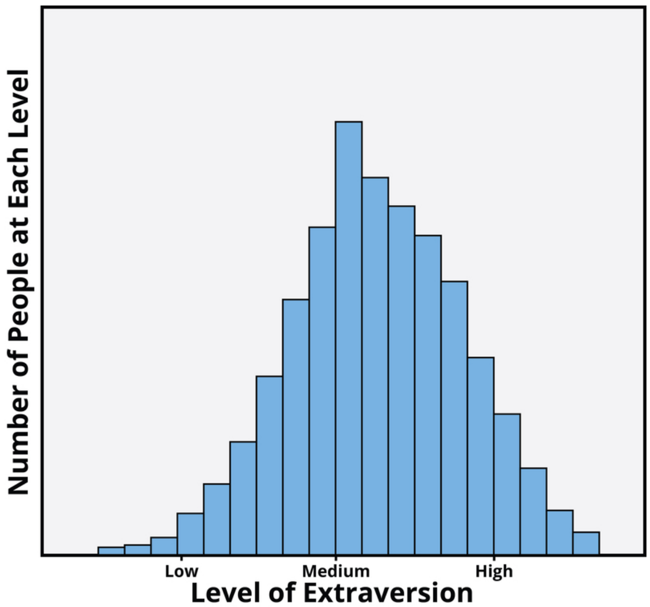


Figure 1. Distribution of Extraversion Scores in a Sample Higher bars mean that more people have scores of that level. This figure shows that most people score towards the middle of the extraversion scale, with fewer people who are highly extraverted or highly introverted.

There are three criteria that are characterize personality traits: (1) consistency, (2) stability, and (3) individual differences.

1. To have a personality trait, individuals must be somewhat consistent across situations in their behaviors related to the trait. For example, if they are talkative at home, they tend also to be talkative at work.

2. Individuals with a trait are also somewhat stable over time in behaviors related to the trait. If they are talkative, for example, at age 30, they will also tend to be talkative at age 40.

3. People differ from one another on behaviors related to the trait. Using speech is not a personality trait and neither is walking on two feet—virtually all individuals do these activities, and there are almost no individual differences. But people differ on how frequently they talk and how active they are, and thus personality traits such as Talkativeness and Activity Level do exist.

A challenge of the trait approach was to discover the major traits on which all people differ. Scientists for many decades generated hundreds of new traits, so that it was soon difficult to keep track and make sense of them. For instance, one psychologist might focus on individual differences in "friendliness," whereas another might focus on the highly related concept of "sociability." Scientists began seeking ways to reduce the number of traits in some systematic way and to discover the basic traits that describe most of the differences between people.

The way that Gordon Allport and his colleague Henry Odbert approached this was to search the dictionary for all descriptors of personality (Allport & Odbert, 1936). Their approach was guided by the **lexical hypothesis**, which states that all important personality characteristics should be reflected in the language that we use to describe other people. Therefore, if we want to understand the fundamental ways in which people differ from one another, we can turn to the words that people use to describe one another. So if we want to know what words people use to describe one another, where should we look? Allport and Odbert looked in the most obvious place—the dictionary. Specifically, they took all the personality descriptors that they could find in the dictionary (they started with almost 18,000 words but quickly reduced that list to a more manageable number) and then used statistical techniques to determine which words “went together.” In other words, if everyone who said that they were “friendly” also said that they were “sociable,” then this might mean that personality psychologists would only need a single trait to capture individual differences in these characteristics. Statistical techniques were used to determine whether a small number of dimensions might underlie all of the thousands of words we use to describe people.

The Five-Factor Model of Personality

Research that used the lexical approach showed that many of the personality descriptors found in the dictionary do indeed overlap. In other words, many of the words that we use to describe people are synonyms. Thus, if we want to know what a person is like, we do not necessarily need to ask how sociable they are, how friendly they are, and how gregarious they are. Instead, because sociable people tend to be friendly and gregarious, we can summarize this personality dimension with a single term. Someone who is sociable, friendly, and gregarious would typically be described as an “Extravert.” Once we know she is an extravert, we can assume that she is sociable, friendly, and gregarious.

Statistical methods (specifically, a technique called **factor analysis**) helped to determine whether a small number of dimensions underlie the diversity of words that people like Allport and Odbert identified. The most widely accepted system to emerge from this approach was “The Big Five” or “**Five-Factor Model**” (Goldberg, 1990; McCrae & John, 1992; McCrae & Costa, 1987). The Big Five comprises five major traits shown in the Figure 2 below. A way to remember these five is with the acronym OCEAN (O is for **Openness**; C is for **Conscientiousness**; E is for **Extraversion**; A is for **Agreeableness**; N is for **Neuroticism**). Figure 3 provides descriptions of people who would score high and low on each of these traits.

Scores on the Big Five traits are mostly independent. That means that a person’s standing on one trait tells very little about their standing on the other traits of the Big Five. For example, a person can be extremely high in Extraversion and be either high or low on Neuroticism.

Big 5 Trait	Definition
Openness	The tendency to appreciate new art, ideas, values, feelings, and behaviors.
Conscientiousness	The tendency to be careful, on-time for appointments, to follow rules, and to be hardworking.
Extraversion	The tendency to be talkative, sociable, and to enjoy others; the tendency to have a dominant style.
Agreeableness	The tendency to agree and go along with others rather than to assert one's own opinions and choices.
Neuroticism	The tendency to frequently experience negative emotions such as anger, worry, and sadness, as well as being interpersonally sensitive.

Figure 2. Descriptions of the Big Five Personality Traits

Similarly, a person can be low in Agreeableness and be either high or low in Conscientiousness. Thus, in the Five-Factor Model, you need five scores to describe most of an individual's personality.

In the Appendix to this module, we present a short scale to assess the Five-Factor Model of personality (Donnellan, Oswald, Baird, & Lucas, 2006). You can take this test to see where you stand in terms of your Big Five scores. John Johnson has also created a helpful website that has personality scales that can be used and taken by the general public:

http://www.personal.psu.edu/j5j/IPIP/ipipneo120.htm

After seeing your scores, you can judge for yourself whether you think such tests are valid.

Traits are important and interesting because they describe stable patterns of behavior that persist for long periods of time (Caspi, Roberts, & Shiner, 2005). Importantly, these stable patterns can have broad-ranging consequences for many areas of our life (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007). For instance, think about the factors that determine success in college. If you were asked to guess what factors predict good grades in college, you might guess something like intelligence. This guess would be correct, but we know much more about who is likely to do well. Specifically, personality researchers have also found the personality traits like Conscientiousness play an important role in college and beyond, probably because highly conscientious individuals study hard, get their work done on time, and are less distracted by nonessential activities that take time away from school work. In addition, highly conscientious people are often healthier than people low in conscientiousness because they are more likely to maintain healthy diets, to exercise, and to follow basic safety procedures like wearing seat belts or bicycle helmets. Over the long term, this consistent pattern of

Big 5 Trait	Example Behavior for LOW Scorers	Example Behavior for HIGH Scorers
Openness	Prefers not to be exposed to alternative moral systems; narrow interests; inartistic; not analytical; down-to-earth	Enjoys seeing people with new types of haircuts and body piercing; curious; imaginative; untraditional
Conscientiousness	Prefers spur-of-the-moment action to planning; unreliable; hedonistic; careless; lax	Never late for a date; organized; hardworking; neat; persevering; punctual; self-disciplined
Extraversion	Preferring a quiet evening reading to a loud party; sober; aloof; unenthusiastic	Being the life of the party; active; optimistic; fun-loving; affectionate
Agreeableness	Quickly and confidently asserts own rights; irritable; manipulative; uncooperative; rude	Agrees with others about political opinions; good-natured; forgiving; gullible; helpful; forgiving
Neuroticism	Not getting irritated by small annoyances; calm, unemotional; hardy; secure; self-satisfied	Constantly worrying about little things; insecure; hypochondriacal; feeling inadequate

Figure 3. Example behaviors for those scoring low and high for the big 5 traits

behaviors can add up to meaningful differences in health and longevity. Thus, personality traits are not just a useful way to describe people you know; they actually help psychologists predict how good a worker someone will be, how long he or she will live, and the types of jobs and activities the person will enjoy. Thus, there is growing interest in personality psychology among psychologists who work in applied settings, such as health psychology or organizational psychology.

Facets of Traits (Subtraits)

So how does it feel to be told that your entire personality can be summarized with scores on just five personality traits? Do you think these five scores capture the complexity of your own and others' characteristic patterns of thoughts, feelings, and behaviors? Most people would probably say no, pointing to some exception in their behavior that goes against the general pattern that others might see. For instance, you may know people who are warm and friendly and find it easy to talk with strangers at a party yet are terrified if they have to perform in front of others or speak to large groups of people. The fact that there are different ways of being extraverted or conscientious shows that there is value in considering lower-level units of personality that are more specific than the Big Five traits. These more specific, lower-level units of personality are often called facets.

To give you a sense of what these narrow units are like, Figure 4 shows facets for each of the

Trait	Facets of Trait
Openness	• Fantasy prone • Open to feelings • Open to diverse behaviors • Open to new and different ideas • Open to various values and beliefs
Conscientiousness	• Competent • Orderly • Dutiful • Achievement oriented • Self-disciplined • Deliberate
Extraversion	• Gregarious (sociable) • Warm • Assertive • Active • Excitement-seeking • Positive emotionality
Agreeableness	• Trusting • Straightforward • Altruistic • Compliant • Modest • Tender-minded
Neuroticism	• Anxious • Angry • Depressed • Self-consciousness • Impulsive • Vulnerable

Figure 4. Facets of Traits

Big Five traits. It is important to note that although personality researchers generally agree about the value of the Big Five traits as a way to summarize one's personality, there is no widely accepted list of facets that should be studied. The list seen here, based on work by researchers Paul Costa and Jeff McCrae, thus reflects just one possible list among many. It should, however, give you an idea of some of the facets making up each of the Five-Factor Model.

Facets can be useful because they provide more specific descriptions of what a person is like. For instance, if we take our friend who loves parties but hates public speaking, we might say that this person scores high on the "gregariousness" and "warmth" facets of extraversion, while scoring lower on facets such as "assertiveness" or "excitement-seeking." This precise profile of facet scores not only provides a better description, it might also allow us to better predict how this friend will do in a variety of different jobs (for example, jobs that require public speaking versus jobs that involve one-on-one interactions with customers; Paunonen & Ashton, 2001). Because different facets within a broad, global trait like extraversion tend to go together (those who are gregarious are often but not always assertive), the broad trait often provides a useful summary of what a person is like. But when we really want to know a person, facet scores add to our knowledge in important ways.

Other Traits Beyond the Five-Factor Model

Despite the popularity of the Five-Factor Model, it is certainly not the only model that exists. Some suggest that there are more than five major traits, or perhaps even fewer. For example, in one of the first comprehensive models to be proposed, Hans Eysenck suggested that

Extraversion and Neuroticism are most important. Eysenck believed that by combining people's standing on these two major traits, we could account for many of the differences in personality that we see in people (Eysenck, 1981). So for instance, a neurotic introvert would be shy and nervous, while a stable introvert might avoid social situations and prefer solitary activities, but he may do so with a calm, steady attitude and little anxiety or emotion. Interestingly, Eysenck attempted to link these two major dimensions to underlying differences in people's biology. For instance, he suggested that introverts experienced too much sensory stimulation and arousal, which made them want to seek out quiet settings and less stimulating environments. More recently, Jeffrey Gray suggested that these two broad traits are related to fundamental reward and avoidance systems in the brain—extraverts might be motivated to seek reward and thus exhibit assertive, reward-seeking behavior, whereas people high in neuroticism might be motivated to avoid punishment and thus may experience anxiety as a result of their heightened awareness of the threats in the world around them (Gray, 1981. This model has since been updated; see Gray & McNaughton, 2000). These early theories have led to a burgeoning interest in identifying the physiological underpinnings of the individual differences that we observe.

Another revision of the Big Five is the **HEXACO model** of traits (Ashton & Lee, 2007). This model is similar to the Big Five, but it posits slightly different versions of some of the traits, and its proponents argue that one important class of individual differences was omitted from the Five-Factor Model. The HEXACO adds Honesty-Humility as a sixth dimension of personality. People high in this trait are sincere, fair, and modest, whereas those low in the trait are manipulative, narcissistic, and self-centered. Thus, trait theorists are agreed that personality traits are important in understanding behavior, but there are still debates on the exact number and composition of the traits that are most important.

There are other important traits that are not included in comprehensive models like the Big Five. Although the five factors capture much that is important about personality, researchers have suggested other traits that capture interesting aspects of our behavior. In Figure 5 below we present just a few, out of hundreds, of the other traits that have been studied by personologists.

Not all of the above traits are currently popular with scientists, yet each of them has experienced popularity in the past. Although the Five-Factor Model has been the target of more rigorous research than some of the traits above, these additional personality characteristics give a good idea of the wide range of behaviors and attitudes that traits can cover.

The Person-Situation Debate and Alternatives to the Trait Perspective

Personality Trait	Description
Machiavellianism	Named after the famous political philosopher, Niccolo Machiavelli, this trait refers to individuals who manipulate the behavior of others, often through duplicity. Machiavellians are often interested in money and power, and pragmatically use others in this quest.
Need for Achievement	Those high in need for achievement want to accomplish a lot and set high standards of excellence for themselves. They are able to work persistently and hard for distant goals. David McClelland argued that economic growth depends in part on citizens with high need for achievement.
Need for Cognition	People high in need for cognition find it rewarding to understand things, and are willing to use considerable cognitive effort in this quest. Such individuals enjoy learning, and the process of trying to understand new things.
Authoritarianism	Authoritarians believe in strict social hierarchies, in which they are totally obedient to those above them, and expect complete obedience from their subordinates. Rigid in adherence to rules, the authoritarian personality is very uncomfortable with uncertainty.
Narcissism	The narcissistic personality has self-love that is so strong that it results in high levels of vanity, conceit, and selfishness. The narcissistic individual often has problems feeling empathetic toward others and grateful to others.
Self-esteem	The tendency to evaluate oneself positively. Self-esteem does not imply that one believes that he or she is better than others, only that he or she is a person of worth.
Optimism	The tendency to expect positive outcomes in the future. People who are optimistic expect good things to happen, and indeed they often have more positive outcomes, perhaps because they work harder to achieve them.
Alexithymia	The inability to recognize and label emotions in oneself. The individual also has a difficult time recognizing emotions in others, and often has difficulties in relationships.

Figure 5. Other Traits Beyond Those Included in the Big Five

The ideas described in this module should probably seem familiar, if not obvious to you. When asked to think about what our friends, enemies, family members, and colleagues are like, some of the first things that come to mind are their personality characteristics. We might think about how warm and helpful our first teacher was, how irresponsible and careless our brother is, or how demanding and insulting our first boss was. Each of these descriptors reflects a personality trait, and most of us generally think that the descriptions that we use for individuals accurately reflect their "characteristic pattern of thoughts, feelings, and behaviors," or in other words, their personality.

The way people behave is only in part a product of their natural personality. Situations also influence how a person behaves. Are you for instance a “different person” as a student in a classroom compared to when you’re a member of a close-knit social group? [Image: UO Education, https://goo.gl/ylgV9T, CC BY-NC 2.0, https://goo.gl/VnKlK8]

But what if this idea were wrong? What if our belief in personality traits were an illusion and people are not consistent from one situation to the next? This was a possibility that shook the foundation of personality psychology in the late 1960s when Walter Mischel published a book called *Personality and Assessment* (1968). In this book, Mischel suggested that if one looks closely at people’s behavior across many different situations, the consistency is really not that impressive. In other words, children who cheat on tests at school may steadfastly follow all rules when playing games and may never tell a lie to their parents. In other words, he suggested, there may not be any general trait of honesty that links these seemingly related behaviors. Furthermore, Mischel suggested that observers may believe that broad personality traits like honesty exist, when in fact, this belief is an illusion. The debate that followed the publication of Mischel’s book was called the **person-situation debate** because it pitted the power of personality against the power of situational factors as determinants of the behavior that people exhibit.

Because of the findings that Mischel emphasized, many psychologists focused on an alternative to the trait perspective. Instead of studying broad, context-free descriptions, like the trait terms we’ve described so far, Mischel thought that psychologists should focus on people’s distinctive reactions to specific situations. For instance, although there may not be a broad and general trait of honesty, some children may be especially likely to cheat on a test when the risk of being caught is low and the rewards for cheating are high. Others might be motivated by the sense of risk involved in cheating and may do so even when the rewards are not very high. Thus, the behavior itself results from the child’s unique evaluation of the risks and rewards present at that moment, along with her evaluation of her abilities and values. Because of this, the same child might act very differently in different situations. Thus, Mischel thought that specific behaviors were driven by the interaction between very specific, psychologically meaningful features of the situation in which people found themselves, the person’s unique way of perceiving that situation, and his or her abilities for dealing with it.

Mischel and others argued that it was these social-cognitive processes that underlie people's reactions to specific situations that provide some consistency when situational features are the same. If so, then studying these broad traits might be more fruitful than cataloging and measuring narrow, context-free traits like Extraversion or Neuroticism.

In the years after the publication of Mischel's (1968) book, debates raged about whether personality truly exists, and if so, how it should be studied. And, as is often the case, it turns out that a more moderate middle ground than what the situationists proposed could be reached. It is certainly true, as Mischel pointed out, that a person's behavior in one specific situation is not a good guide to how that person will behave in a very different specific situation. Someone who is extremely talkative at one specific party may sometimes be reticent to speak up during class and may even act like a wallflower at a different party. But this does not mean that personality does not exist, nor does it mean that people's behavior is completely determined by situational factors. Indeed, research conducted after the person-situation debate shows that on average, the effect of the "situation" is about as large as that of personality traits. However, it is also true that if psychologists assess a broad range of behaviors across many different situations, there are general tendencies that emerge. Personality traits give an indication about how people will act on average, but frequently they are not so good at predicting how a person will act in a specific situation at a certain moment in time. Thus, to best capture broad traits, one must assess *aggregate* behaviors, averaged over time and across many different types of situations. Most modern personality researchers agree that there is a place for broad personality traits and for the narrower units such as those studied by Walter Mischel.

Appendix

The Mini-IPIP Scale

(Donnellan, Oswald, Baird, & Lucas, 2006)

Instructions: Below are phrases describing people's behaviors. Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. Please read each statement carefully, and put a number from 1 to 5 next to it to describe how accurately the statement describes you.

1 = Very inaccurate

2 = Moderately inaccurate

3 = Neither inaccurate nor accurate

4 = Moderately accurate

5 = Very accurate

1. _______ Am the life of the party (E)
2. _______ Sympathize with others' feelings (A)
3. _______ Get chores done right away (C)
4. _______ Have frequent mood swings (N)
5. _______ Have a vivid imagination (O)
6. _______Don't talk a lot (E)
7. _______ Am not interested in other people's problems (A)
8. _______ Often forget to put things back in their proper place (C)
9. _______ Am relaxed most of the time (N)
10. ______ Am not interested in abstract ideas (O)
11. ______ Talk to a lot of different people at parties (E)
12. ______ Feel others' emotions (A)
13. ______ Like order (C)
14. ______ Get upset easily (N)
15. ______ Have difficulty understanding abstract ideas (O)
16. ______ Keep in the background (E)
17. ______ Am not really interested in others (A)
18. ______ Make a mess of things (C)
19. ______ Seldom feel blue (N)
20. ______ Do not have a good imagination (O)

Scoring: The first thing you must do is to reverse the items that are worded in the opposite direction. In order to do this, subtract the number you put for that item from 6. So if you put a 4, for instance, it will become a 2. Cross out the score you put when you took the scale, and put the new number in representing your score subtracted from the number 6.

Items to be reversed in this way: 6, 7, 8, 9, 10, 15, 16, 17, 18, 19, 20

Next, you need to add up the scores for each of the five OCEAN scales (including the reversed numbers where relevant). Each OCEAN score will be the sum of four items. Place the sum next to each scale below.

__________ Openness: Add items 5, 10, 15, 20

__________ Conscientiousness: Add items 3, 8, 13, 18

__________ Extraversion: Add items 1, 6, 11, 16

__________ Agreeableness: Add items 2, 7, 12, 17

__________ Neuroticism: Add items 4, 9,14, 19

Compare your scores to the norms below to see where you stand on each scale. If you are low on a trait, it means you are the opposite of the trait label. For example, low on Extraversion is Introversion, low on Openness is Conventional, and low on Agreeableness is Assertive.

19–20 Extremely High, 17–18 Very High, 14–16 High,

11–13 Neither high nor low; in the middle, 8–10 Low, 6–7 Very low, 4–5 Extremely low

Outside Resources

Video 1: Gabriela Cintron's – 5 Factors of Personality (OCEAN Song). This is a student-made video which cleverly describes, through song, common behavioral characteristics of the Big 5 personality traits. It was one of the winning entries in the 2016-17 Noba + Psi Chi Student Video Award.
https://www.youtube.com/watch?feature=youtu.be&v=Rk8CDXMb8_U&app=desktop

Video 2: Michael Harris' – Personality Traits: The Big 5 and More. This is a student-made video that looks at characteristics of the OCEAN traits through a series of funny vignettes. It also presents on the Person vs Situation Debate. It was one of the winning entries in the 2016-17 Noba + Psi Chi Student Video Award.
https://vimeo.com/218245492

Video 3: David M. Cole's – Grouchy with a Chance of Stomping. This is a student-made video that makes a very important point about the relationship between personality traits and behavior using a handy weather analogy. It was one of the winning entries in the 2016-17 Noba + Psi Chi Student Video Award.
https://www.youtube.com/watch?v=GnaFMjaJtlY

Web: International Personality Item Pool
http://ipip.ori.org/

Web: John Johnson personality scales
http://www.personal.psu.edu/j5j/IPIP/ipipneo120.htm

Web: Personality trait systems compared
http://www.personalityresearch.org/bigfive/goldberg.html

Web: Sam Gosling website
http://homepage.psy.utexas.edu/homepage/faculty/gosling/samgosling.htm

Discussion Questions

1. Consider different combinations of the Big Five, such as O (Low), C (High), E (Low), A (High), and N (Low). What would this person be like? Do you know anyone who is like this? Can you select politicians, movie stars, and other famous people and rate them on the Big Five?

2. How do you think learning and inherited personality traits get combined in adult personality?
3. Can you think of instances where people do not act consistently—where their personality traits are not good predictors of their behavior?
4. Has your personality changed over time, and in what ways?
5. Can you think of a personality trait not mentioned in this module that describes how people differ from one another?
6. When do extremes in personality traits become harmful, and when are they unusual but productive of good outcomes?

Vocabulary

Agreeableness
A personality trait that reflects a person's tendency to be compassionate, cooperative, warm, and caring to others. People low in agreeableness tend to be rude, hostile, and to pursue their own interests over those of others.

Conscientiousness
A personality trait that reflects a person's tendency to be careful, organized, hardworking, and to follow rules.

Continuous distributions
Characteristics can go from low to high, with all different intermediate values possible. One does not simply have the trait or not have it, but can possess varying amounts of it.

Extraversion
A personality trait that reflects a person's tendency to be sociable, outgoing, active, and assertive.

Facets
Broad personality traits can be broken down into narrower facets or aspects of the trait. For example, extraversion has several facets, such as sociability, dominance, risk-taking and so forth.

Factor analysis
A statistical technique for grouping similar things together according to how highly they are associated.

Five-Factor Model
(also called the Big Five) The Five-Factor Model is a widely accepted model of personality traits. Advocates of the model believe that much of the variability in people's thoughts, feelings, and behaviors can be summarized with five broad traits. These five traits are Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.

HEXACO model
The HEXACO model is an alternative to the Five-Factor Model. The HEXACO model includes six traits, five of which are variants of the traits included in the Big Five (Emotionality [E], Extraversion [X], Agreeableness [A], Conscientiousness [C], and Openness [O]). The sixth

factor, Honesty-Humility [H], is unique to this model.

Independent
Two characteristics or traits are separate from one another-- a person can be high on one and low on the other, or vice-versa. Some correlated traits are relatively independent in that although there is a tendency for a person high on one to also be high on the other, this is not always the case.

Lexical hypothesis
The lexical hypothesis is the idea that the most important differences between people will be encoded in the language that we use to describe people. Therefore, if we want to know which personality traits are most important, we can look to the language that people use to describe themselves and others.

Neuroticism
A personality trait that reflects the tendency to be interpersonally sensitive and the tendency to experience negative emotions like anxiety, fear, sadness, and anger.

Openness to Experience
A personality trait that reflects a person's tendency to seek out and to appreciate new things, including thoughts, feelings, values, and experiences.

Personality
Enduring predispositions that characterize a person, such as styles of thought, feelings and behavior.

Personality traits
Enduring dispositions in behavior that show differences across individuals, and which tend to characterize the person across varying types of situations.

Person-situation debate
The person-situation debate is a historical debate about the relative power of personality traits as compared to situational influences on behavior. The situationist critique, which started the person-situation debate, suggested that people overestimate the extent to which personality traits are consistent across situations.

References

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycholexical study. *Psychological Monographs, 47*, 211.

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychological Review, 11*, 150–166.

Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annual Reviews of Psychology, 56*, 453–484.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*, 192–203.

Eysenck, H. J. (1981). *A model for personality*.New York: Springer Verlag.

Goldberg, L. R. (1990). An alternative description of personality: The Big Five personality traits. *Journal of Personality and Social Psychology, 59*, 1216–1229.

Gray, J. A. (1981). A critique of Eysenck's theory of personality. In H. J. Eysenck (Ed.), *A Model for Personality* (pp. 246-276). New York: Springer Verlag.

Gray, J. A. & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system (second edition)*.Oxford: Oxford University Press.

Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality traits*. Cambridge, UK: Cambridge University Press.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81–90.

McCrae, R. R. & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60*, 175–215.

Mischel, W. (1968). *Personality and assessment*. New York: John Wiley.

Paunonen, S. V., & Ashton, M. S. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology, 81*, 524–539.

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Golberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*, 313-345.

Emotions and Motivation

33
Affective Neuroscience

Eddie Harmon-Jones & Cindy Harmon-Jones

This module provides a brief overview of the neuroscience of emotion. It integrates findings from human and animal research to describe the brain networks and associated neurotransmitters involved in basic affective systems.

Learning Objectives

- Define affective neuroscience.
- Describe neuroscience techniques used to study emotions in humans and animals.
- Name five emotional systems and their associated neural structures and neurotransmitters.
- Give examples of exogenous chemicals (e.g., drugs) that influence affective systems, and discuss their effects.
- Discuss multiple affective functions of the amygdala and the nucleus accumbens.
- Name several specific human emotions, and discuss their relationship to the affective systems of nonhuman animals.

Affective Neuroscience: What is it?

Affective neuroscience examines how the brain creates emotional responses. Emotions are psychological phenomena that involve changes to the body (e.g., facial expression), changes in autonomic nervous system activity, feeling states (subjective responses), and urges to act in specific ways (motivations; Izard, 2010). Affective **neuroscience** aims to understand how matter (brain structures and chemicals) creates one of the most fascinating aspects of mind, the emotions. Affective neuroscience uses unbiased, observable measures that provide credible evidence to other sciences and laypersons on the importance of emotions. It also leads to biologically based treatments for affective disorders (e.g., depression).

Although we all experience emotions all the time, they are very difficult to describe and study. Fortunately, technological advances are making this easier. [Image: Waag Society, https://goo.gl/F0KdnB, CC BY-NC-SA 2.0, https://goo.gl/iF4hmM]

The human brain and its responses, including emotions, are complex and flexible. In comparison, nonhuman animals possess simpler nervous systems and more basic emotional responses. Invasive neuroscience techniques, such as electrode implantation, lesioning, and hormone administration, can be more easily used in animals than in humans. Human neuroscience must rely primarily on noninvasive techniques such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), and on studies of individuals with brain lesions caused by accident or disease. Thus, animal research provides useful models for understanding affective processes in humans. Affective circuits found in other species, particularly social mammals such as rats, dogs, and monkeys, function similarly to human affective networks, although nonhuman animals' brains are more basic.

In humans, emotions and their associated neural systems have additional layers of complexity and flexibility. Compared to animals, humans experience a vast variety of nuanced and sometimes conflicting emotions. Humans also respond to these emotions in complex ways, such that conscious goals, values, and other cognitions influence behavior in addition to emotional responses. However, in this module we focus on the similarities between organisms, rather than the differences. We often use the term "organism" to refer to the individual who is experiencing an emotion or showing evidence of particular neural activations. An organism could be a rat, a monkey, or a human.

Across species, emotional responses are organized around the organism's survival and reproductive needs. Emotions influence perception, cognition, and behavior to help organisms survive and thrive (Farb, Chapman, & Anderson, 2013). Networks of structures in the brain respond to different needs, with some overlap between different emotions. Specific emotions are not located in a single structure of the brain. Instead, emotional responses involve networks of activation, with many parts of the brain activated during any emotional process. In fact, the brain circuits involved in emotional reactions include nearly the entire brain (Berridge & Kringelbach, 2013). Brain circuits located deep within the brain below the cerebral cortex are primarily responsible for generating basic emotions (Berridge & Kringelbach, 2013; Panksepp & Biven, 2012). In the past, research attention was focused on specific brain structures that will be reviewed here, but future research may find that additional areas of the brain are also important in these processes.

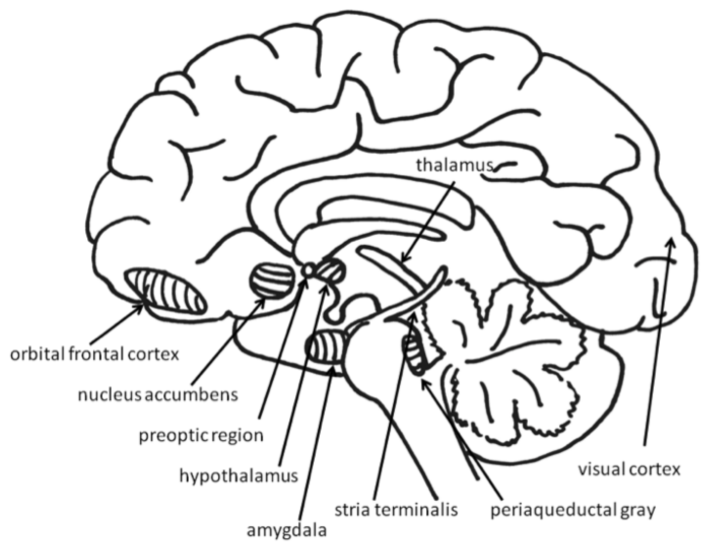


Figure 1: Structures in the brain

Basic Emotions

Desire: The neural systems of reward seeking

One of the most important affective neuronal systems relates to feelings of desire, or the appetite for rewards. Researchers refer to these appetitive processes using terms such as "wanting" (Berridge & Kringelbach, 2008), "seeking" (Panksepp & Biven, 2012), or "behavioural activation sensitivity" (Gray, 1987). When the appetitive system is aroused, the organism shows enthusiasm, interest, and curiosity. These neural circuits motivate the animal to move through its environment in search of rewards such as appetizing foods, attractive sex partners, and other pleasurable stimuli. When the appetitive system is underaroused, the organism appears depressed and helpless.

Much evidence for the structures involved in this system comes from animal research using direct brain stimulation. When an electrode is implanted in the lateral hypothalamus or in cortical or mesencephalic regions to which the hypothalamus is connected, animals will press a lever to deliver electrical stimulation, suggesting that they find the stimulation pleasurable. The regions in the desire system also include the amygdala, nucleus accumbens, and frontal cortex (Panksepp & Biven, 2012). The neurotransmitter dopamine, produced in the mesolimbic and mesocortical dopamine circuits, activates these regions. It creates a sense of excitement, meaningfulness, and anticipation. These structures are also sensitive to drugs such as cocaine and amphetamines, chemicals that have similar effects to dopamine (Panksepp & Biven, 2012).

Research in both humans and nonhuman animals shows that the left frontal cortex (compared to the right frontal cortex) is more active during appetitive emotions such as desire and interest. Researchers first noted that persons who had suffered damage to the left frontal cortex developed depression, whereas those with damage to the right frontal cortex developed mania (Goldstein, 1939). The relationship between left frontal activation and approach-related emotions has been confirmed in healthy individuals using EEG and fMRI (Berkman & Lieberman, 2010). For example, increased left frontal activation occurs in 2- to 3-day-old infants when sucrose is placed on their tongues (Fox & Davidson, 1986), and in hungry adults as they view pictures of desirable desserts (Gable & Harmon-Jones, 2008). In addition, greater left frontal activity in appetitive situations has been found to relate to dopamine (Wacker, Mueller, Pizzagalli, Hennig, & Stemmler, 2013).

Just looking at an image of appealing food should increase the activity in your left frontal cortex. Yum! [Image: CC0 Public Domain, https://goo.gl/m25gce]

"Liking" : The neural circuits of pleasure and enjoyment

Surprisingly, the amount of desire an individual feels toward a reward need not correspond to how much he or she likes that reward. This is because the neural structures involved in the enjoyment of rewards are different from the structures involved in the desire for the rewards. “Liking” (e.g., enjoyment of a sweet liquid) can be measured in babies and nonhuman animals by measuring licking speed, tongue protrusions, and happy facial expressions, whereas “wanting” (desire) is shown by the willingness to work hard to obtain a reward (Berridge & Kringelbach, 2008). Liking has been distinguished from wanting in research on topics such as drug abuse. For example, drug addicts often desire drugs even when they know that the ones available will not provide pleasure (Stewart, de Wit, & Eikelboom, 1984).

Research on liking has focused on a small area within the nucleus accumbens and on the posterior half of the ventral pallidum. These brain regions are sensitive to opioids and endocannabinoids. Stimulation of other regions of the reward system increases wanting, but does not increase liking, and in some cases even decreases liking. The research on the distinction between desire and enjoyment contributes to the understanding of human addiction, particularly why individuals often continue to frantically pursue rewards such as cocaine, opiates, gambling, or sex, even when they no longer experience pleasure from obtaining these rewards due to habituation.

The experience of pleasure also involves the orbitofrontal cortex. Neurons in this region fire when monkeys taste, or merely see pictures of, desirable foods. In humans, this region is activated by pleasant stimuli including money, pleasant smells, and attractive faces (Gottfried, O’Doherty & Dolan, 2002; O’Doherty, Deichmann, Critchley, & Dolan, 2002; O’Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; O’Doherty, Winston, Critchley, Perrett, Burt, & Dolan, 2003).

Because fear is so important for our survival (i.e., fear informs us when something threatens us), our brains are able to “recognize” frightening stimuli before we ourselves are even consciously aware of them. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Fear: The neural system of freezing and fleeing

Fear is an unpleasant emotion that motivates avoidance of potentially harmful situations. Slight stimulation of the fear-related areas in the brain causes animals to freeze, whereas intense stimulation causes them to flee. The fear circuit extends from the central amygdala to the periaqueductal gray in the midbrain. These structures are sensitive to glutamate, corticotrophin releasing factor, adreno-cortico-trophic hormone, cholecystokinin, and several different neuropeptides. Benzodiazepines and other tranquilizers inhibit activation in these areas (Panksepp & Biven, 2012).

The role of the amygdala in fear responses has been extensively studied. Perhaps because fear is so important to survival, two pathways send signals to the amygdala from the sensory organs. When an individual sees a snake, for example, the sensory information travels from the eye to the thalamus and then to the visual cortex. The visual cortex sends the information on to the amygdala, provoking a fear response. However, the thalamus also quickly sends the information straight to the amygdala, so that the organism can react before consciously perceiving the snake (LeDoux, Farb, & Ruggiero, 1990). The pathway from the thalamus to the amygdala is fast but less accurate than the slower pathway from the visual cortex. Damage to the amygdala or areas of the ventral hypocampus interferes with fear conditioning in both humans and nonhuman animals (LeDoux, 1996).

Rage: The circuits of anger and attack

Anger or rage is an arousing, unpleasant emotion that motivates organisms to approach and attack (Harmon-Jones, Harmon-Jones, & Price, 2013). Anger can be evoked through goal frustration, physical pain, or physical restraint. In territorial animals, anger is provoked by a stranger entering the organism's home territory (Blanchard & Blanchard, 2003). The neural networks for anger and fear are near one another, but separate (Panksepp & Biven, 2012). They extend from the medial amygdala, through specific parts of the hypothalamus, and into the periaqueductal gray of the midbrain. The anger circuits are linked to the appetitive circuits, such that lack of an anticipated reward can provoke rage. In addition, when humans are angered, they show increased left frontal cortical activation, supporting the idea that anger is an approach-related emotion (Harmon-Jones et al., 2013). The neurotransmitters involved in rage are not yet well understood, but Substance P may play an important role (Panksepp & Biven, 2012). Other neurochemicals that may be involved in anger include testosterone (Peterson & Harmon-Jones, 2012) and arginine-vasopressin (Heinrichs, von Dawans, & Domes, 2009). Several chemicals inhibit the rage system, including opioids and high doses of antipsychotics, such as chlorpromazine (Panksepp & Biven, 2012).

Love: The neural systems of care and attachment

Just as scientists today distinguish between types of love like "romantic" and "parental," so did the ancient Greeks who used the terms "eros" and "storge." [Image: Go-tea 郭天, https://goo.gl/KLFxJP, CC BY 2.0, https://goo.gl/zHmGV2]

For social animals such as humans, attachment to other members of the same species produces the positive emotions of attachment: love, warm feelings, and affection. The emotions that motivate nurturing behavior (e.g., maternal care) are distinguishable from those that motivate staying close to an attachment figure in order to receive care and protection (e.g., infant attachment). Important regions for maternal nurturing include the dorsal preoptic area (Numan & Insel, 2003) and the bed nucleus of the stria terminalis (Panksepp, 1998). These regions overlap with the areas involved in sexual desire, and are sensitive to some of the same neurotransmitters, including oxytocin, arginine-vasopressin, and endogenous opioids (endorphins and enkephalins).

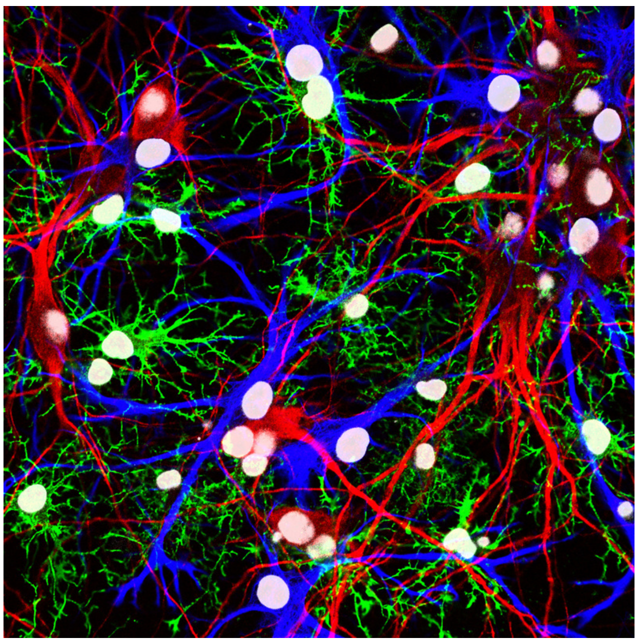
Grief: The neural networks of loneliness and panic

The neural networks involved in infant attachment are also sensitive to separation. These regions produce the painful emotions of grief, panic, and loneliness. When infant humans or other infant mammals are separated from their mothers, they produce distress vocalizations, or crying. The attachment circuits are those that cause organisms to produce distress vocalizations when electrically stimulated.

The attachment system begins in the midbrain periaqueductal gray, very close to the area that produces physical pain responses, suggesting that it may have originated from the pain circuits (Panksepp, 1998). Separation distress can also be evoked by stimulating the dorsomedial thalamus, ventral septum, dorsal preoptic region, and areas in the bed nucleus of stria terminalis (near sexual and maternal circuits; Panksepp, Normansell, Herman, Bishop, & Crepeau, 1988).

These regions are sensitive to endogenous opiates, oxytocin, and prolactin. All of these neurotransmitters *prevent* separation distress. Opiate drugs such as morphine and heroin, as well as nicotine, artificially produce feelings of pleasure and gratification, similar to those normally produced during positive social interactions. This may explain why these drugs are addictive. Panic attacks appear to be an intense form of separation distress triggered by the attachment system, and panic can be effectively relieved by opiates. Testosterone also reduces separation distress, perhaps by reducing attachment needs. Consistent with this, panic attacks are more common in women than in men.

Plasticity: Experiences can alter the brain



Neural plasticity can be summed up in the phrase: "Neurons that fire together, wire together." Or in other words, when certain emotions are paired with certain contexts, we learn to associate the two together. [Image: NICHD NIH, https://goo.gl/T3PUlg, CC BY 2.0, https://goo.gl/BRvSA7]

The responses of specific neural regions may be modified by experience. For example, the front shell of the nucleus accumbens is generally involved in appetitive behaviors, such as eating, and the back shell is generally involved in fearful defensive behaviors (Reynolds & Berridge, 2001, 2002). Research using human neuroimaging has also revealed this front–back distinction in the functions of the nucleus accumbens (Seymour, Daw, Dayan, Singer, & Dolan, 2007). However, when rats are exposed to stressful environments, their fear-generating regions expand toward the front, filling almost 90% of the nucleus accumbens shell. On the other hand, when rats are exposed to preferred home environments, their fear-generating regions shrink and the appetitive regions expand toward the back, filling approximately 90% of the shell (Reynolds & Berridge, 2008).

Brain structures have multiple functions

Although much affective neuroscience research has emphasized whole structures, such as the amygdala and nucleus accumbens, it is important to note that many of these structures are more accurately referred to as complexes. They include distinct groups of nuclei that perform different tasks. At present, human neuroimaging techniques such as fMRI are unable to examine the activity of individual nuclei in the way that invasive animal neuroscience can.

For instance, the amygdala of the nonhuman primate can be divided into 13 nuclei and cortical areas (Freese & Amaral, 2009). These regions of the amygdala perform different functions. The central nucleus sends outputs involving brainstem areas that result in innate emotional expressions and associated physiological responses. The basal nucleus is connected with striatal areas that are involved with actions such as running toward safety. Furthermore, it is not possible to make one-to-one maps of emotions onto brain regions. For example, extensive research has examined the involvement of the amygdala in fear, but research has also shown that the amygdala is active during uncertainty (Whalen, 1998) as well as positive emotions (Anderson et al., 2003; Schulkin, 1990).

Conclusion

Research in affective neuroscience has contributed to knowledge regarding emotional, motivational, and behavioral processes. The study of the basic emotional systems of nonhuman animals provides information about the organization and development of more complex human emotions. Although much still remains to be discovered, current findings in affective neuroscience have already influenced our understanding of drug use and abuse, psychological disorders such as panic disorder, and complex human emotions such as desire and enjoyment, grief and love.

Outside Resources

Video: A 1-hour interview with Jaak Panksepp, the father of affective neuroscience
http://www.youtube.com/watch?v=u4ICY6-7hJo

Video: A 15-minute interview with Kent Berridge on pleasure in the brain
http://www.youtube.com/watch?v=51rGE1DgIo0

Video: A 5-minute interview with Joseph LeDoux on the amygdala and fear
http://www.youtube.com/watch?v=fDD5wvFMH6U

Web: Brain anatomy interactive 3D model
http://www.pbs.org/wnet/brain/3d/index.html

Discussion Questions

1. The neural circuits of “liking” are different from the circuits of “wanting.” How might this relate to the problems people encounter when they diet, fight addictions, or try to change other habits?
2. The structures and neurotransmitters that produce pleasure during social contact also produce panic and grief when organisms are deprived of social contact. How does this contribute to an understanding of love?
3. Research shows that stressful environments increase the area of the nucleus accumbens that is sensitive to fear, whereas preferred environments increase the area that is sensitive to rewards. How might these changes be adaptive?

Vocabulary

Affect
An emotional process; includes moods, subjective feelings, and discrete emotions.

Amygdala
Two almond-shaped structures located in the medial temporal lobes of the brain.

Hypothalamus
A brain structure located below the thalamus and above the brain stem.

Neuroscience
The study of the nervous system.

Nucleus accumbens
A region of the basal forebrain located in front of the preoptic region.

Orbital frontal cortex
A region of the frontal lobes of the brain above the eye sockets.

Periaqueductal gray
The gray matter in the midbrain near the cerebral aqueduct.

Preoptic region
A part of the anterior hypothalamus.

Stria terminalis
A band of fibers that runs along the top surface of the thalamus.

Thalamus
A structure in the midline of the brain located between the midbrain and the cerebral cortex.

Visual cortex
The part of the brain that processes visual information, located in the back of the brain.

References

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., . . . Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience, 6*, 196–202.

Berkman, E. T., & Lieberman, M. D. (2010). Approaching the bad and avoiding the good: Lateral prefrontal cortical asymmetry distinguishes between action and valence. *Journal of Cognitive Neuroscience, 22*(9), 1970–1979. doi: 10.1162/jocn.2009.21317

Berridge, K. C., & Kringelbach, M. L. (2013). Neuroscience of affect: brain mechanisms of pleasure and displeasure. *Current Opinion in Neurobiology, 23*, 294–303. doi.org/10.1016/j.conb.2013.01.017

Berridge, K. C., & Kringelbach, M. L. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology, 199*, 457–480. doi: 10.1007/s00213-008-1099-6

Blanchard, D. C., & Blanchard, R. J. (2003). What can animal aggression research tell us about human aggression? *Hormones and Behavior, 44*, 171–177.

Farb, N.A.S., Chapman, H. A., & Anderson, A. K. (2013). Emotions: Form follows function. *Current Opinion in Neurobiology, 23*, 393–398. http://dx.doi.org/10.1016/j.conb.2013.01.015

Fox, N. A., & Davidson, R. J. (1986). Taste-elicited changes in facial signs of emotion and the asymmetry of brain electrical activity in human newborns. *Neuropsychologia, 24*, 417–422.

Freese, J. L., & Amaral, D. G. (2009). Neuroanatomy of the primate amygdala. In P. J. Whalen & E. A. Phelps (Eds.), *The human amygdala* (pp. 3–42). New York, NY: Guilford Press.

Gable, P. A., & Harmon-Jones, E. (2008). Relative left frontal activation to appetitive stimuli: Considering the role of individual differences. *Psychophysiology, 45*, 275-278.

Goldstein, K. (1939). *The organism: An holistic approach to biology, derived from pathological data in man*. New York, NY: American Book.

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *Journal of Neuroscience, 22*, 10829–10837.

Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge, England: Cambridge University Press.

Harmon-Jones, E., Harmon-Jones, C., & Price, T. F. (2013). What is approach motivation? *Emotion Review, 5*, 291–295. doi: 10.1177/1754073913477509

Heinrichs, M., von Dawans, B., & Domes, G. (2009). Oxytocin, vasopressin, and human social behavior. *Frontiers in Neuroendocrinology, 30*, 548–557.

Izard, C. E. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review, 2*, 363–370. doi: 10.1177/1754073910374661

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life*. New York, NY: Simon & Schuster.

LeDoux, J. E., Farb, C. F., Ruggiero, D. A. (1990). Topographic organization of neurons in the acoustic thalamus that project to the amygdala. *Journal of Neuroscience, 10*, 1043–1054.

Numan, M., & Insel, T. R. (2003). *The neurobiology of parental behavior*. New York, NY: SpringerVerlag.

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.

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*, 95–102.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*, 147–155.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York, NY: Oxford University Press.

Panksepp, J., & Biven, L. (2012). *The archaeology of mind: Neuroevolutionary origins of human emotions*. New York, NY: Norton.

Panksepp, J., Normansell, L., Herman, B., Bishop, P., & Crepeau, L. (1988). Neural and neurochemical control of the separation distress call. In J. D. Newman (Ed.), *The physiological control of mammalian vocalization* (pp. 263–299). New York, NY: Plenum.

Peterson, C. K., & Harmon-Jones, E. (2012). Anger and testosterone: Evidence that situationally-induced anger relates to situationally-induced testosterone. *Emotion, 12*, 899–902. doi: 10.1037/a0025300

Reynolds, S. M., & Berridge, K. C. (2008). Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nature Neuroscience, 11*, 423–425.

Reynolds, S. M., & Berridge, K. C. (2002). Positive and negative motivation in nucleus accumbens shell: Bivalent rostrocaudal gradients for GABA-elicited eating, taste "liking"/"disliking" reactions, place preference/avoidance, and fear. *Journal of Neuroscience, 22*, 7308–7320.

Reynolds, S. M., & Berridge, K. C. (2001). Fear and feeding in the nucleus accumbens shell: Rostrocaudal segregation of GABA-elicited defensive behavior versus eating behavior.

Journal of Neuroscience, 21, 3261–3270.

Schulkin, J. (1991). *Sodium hunger: The search for a salty taste*. New York, NY: Cambridge University Press.

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience, 27*, 4826–4831.

Stewart, J., De Wit, H., & Eikelboom, R. (1984). Role of unconditioned and conditioned drug effects in the self-administration of opiates and stimulants. *Psychological Review, 91*, 251-268.

Wacker, J., Mueller, E. M., Pizzagalli, D. A., Hennig, J., & Stemmler, G. (2013). Dopamine-D2-receptor blockade reverses the association between trait approach motivation and frontal asymmetry in an approach-motivation context. *Psychological Science, 24*(4), 489–497. doi: 10.1177/0956797612458935

Whalen, P. J. (1998). Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science, 7*, 177–188.

34
Functions of Emotions

Hyisung Hwang & David Matsumoto

Emotions play a crucial role in our lives because they have important functions. This module describes those functions, dividing the discussion into three areas: the intrapersonal, the interpersonal, and the social and cultural functions of emotions. The section on the intrapersonal functions of emotion describes the roles that emotions play within each of us individually; the section on the interpersonal functions of emotion describes the meanings of emotions to our relationships with others; and the section on the social and cultural functions of emotion describes the roles and meanings that emotions have to the maintenance and effective functioning of our societies and cultures at large. All in all we will see that emotions are a crucially important aspect of our psychological composition, having meaning and function to each of us individually, to our relationships with others in groups, and to our societies as a whole.

Learning Objectives

- Gain an appreciation of the importance of emotion in human life.
- Understand the functions and meanings of emotion in three areas of life: the intrapersonal, interpersonal, and social–cultural.
- Give examples of the role and function of emotion in each of the three areas described.

Introduction

It is impossible to imagine life without emotion. We treasure our feelings—the joy at a ball game, the pleasure of the touch of a loved one, or the fun with friends on a night out. Even negative emotions are important, such as the sadness when a loved one dies, the anger when violated, the fear that overcomes us in a scary or unknown situation, or the guilt or shame toward others when our sins are made public. Emotions color life experiences and give those experiences meaning and flavor.

In fact, emotions play many important roles in people's lives and have been the topic of scientific inquiry in psychology for well over a century (Cannon, 1927; Darwin, 1872; James, 1890). This module explores why we have emotions and why they are important. Doing so requires us to understand the function of emotions, and this module does so below by dividing the discussion into three sections. The first concerns the intrapersonal functions of emotion, which refer to the role that emotions play within each of us individually. The second concerns the interpersonal functions of emotion, which refer to the role emotions play between individuals within a group. The third concerns the social and cultural functions of emotion, which refer to the role that emotions play in the maintenance of social order within a society. All in all, we will see that emotions inform us of who we are, what our relationships with others are like, and how to behave in social interactions. Emotions give meaning to events; without emotions, those events would be mere facts. Emotions help coordinate interpersonal relationships. And emotions play an important role in the cultural functioning of keeping human societies together.

Emotions help us navigate the complex social landscape of our lives. [Image: Gwenaël Piase, https://goo.gl/d4EDKS, CC BY-NC-SA 2.0, https://goo.gl/hSpkVI]

Intrapersonal Functions of Emotion

Emotions Help us Act Quickly with Minimal Conscious Awareness

Emotions are rapid information-processing systems that help us act with minimal thinking (Tooby & Cosmides, 2008). Problems associated with birth, battle, death, and seduction have occurred throughout evolutionary history and emotions evolved to aid humans in adapting to those problems rapidly and with minimal conscious cognitive intervention. If we did not have emotions, we could not make rapid decisions concerning whether to attack, defend, flee, care for others, reject food, or approach something useful, all of which were functionally adaptive in our evolutionary history and helped us to survive. For instance, drinking spoiled milk or eating rotten eggs has negative consequences for our welfare. The emotion of disgust, however, helps us immediately take action by not ingesting them in the first place or by vomiting them out. This response is adaptive because it aids, ultimately, in our survival and allows us to act immediately without much thinking. In some instances, taking the time to sit and rationally think about what to do, calculating cost–benefit ratios in one's mind, is a luxury that might cost one one's life. Emotions evolved so that we can act without that depth of thinking.

Emotions Prepare the Body for Immediate Action

The emotion of disgust serves to protect us from toxins and contamination, of the physical and moral variety. [Image: Runs with Scissors, https://goo.gl/FQRxGa, CC BY-NC 2.0, https://goo.gl/tgFydH]

Emotions prepare us for behavior. When triggered, emotions orchestrate systems such as perception, attention, inference, learning, memory, goal choice, motivational priorities, physiological reactions, motor behaviors, and behavioral decision making (Cosmides & Tooby, 2000; Tooby & Cosmides, 2008). Emotions simultaneously activate certain systems and deactivate others in order to prevent the chaos of competing systems operating at the same time, allowing for coordinated responses to environmental stimuli (Levenson, 1999). For instance, when we are afraid, our bodies shut down temporarily unneeded digestive processes, resulting in saliva reduction (a dry mouth); blood flows disproportionately to the lower half of the body; the visual field expands; and air is breathed in, all preparing the body to flee. Emotions initiate a system of components that includes subjective experience, expressive behaviors, physiological reactions, action tendencies, and cognition, all for the purposes of specific actions; the term "emotion" is, in reality, a metaphor for these reactions.

One common misunderstanding many people have when thinking about emotions, however, is the belief that emotions must always directly produce action. This is not true. Emotion certainly *prepares* the body for action; but whether people actually engage in action is dependent on many factors, such as the context within which the emotion has occurred, the target of the emotion, the perceived consequences of one's actions, previous experiences, and so forth (Baumeister, Vohs, DeWall, & Zhang, 2007; Matsumoto & Wilson, 2008). Thus, emotions are just one of many determinants of behavior, albeit an important one.

Emotions Influence Thoughts

Emotions are also connected to thoughts and memories. Memories are not just facts that are encoded in our brains; they are colored with the emotions felt at those times the facts occurred (Wang & Ross, 2007). Thus, emotions serve as the neural glue that connects those disparate facts in our minds. That is why it is easier to remember happy thoughts when happy, and angry times when angry. Emotions serve as the affective basis of many attitudes, values, and beliefs that we have about the world and the people around us; without emotions those attitudes, values, and beliefs would be just statements without meaning, and emotions give those statements meaning. Emotions influence our thinking processes, sometimes in constructive ways, sometimes not. It is difficult to think critically and clearly when we feel intense emotions, but easier when we are not overwhelmed with emotions (Matsumoto, Hirayama, & LeRoux, 2006).

Emotions Motivate Future Behaviors

Because emotions prepare our bodies for immediate action, influence thoughts, and can be felt, they are important motivators of future behavior. Many of us strive to experience the feelings of satisfaction, joy, pride, or triumph in our accomplishments and achievements. At the same time, we also work very hard to avoid strong negative feelings; for example, once we have felt the emotion of disgust when drinking the spoiled milk, we generally work very hard to avoid having those feelings again (e.g., checking the expiration date on the label before buying the milk, smelling the milk before drinking it, watching if the milk curdles in one's coffee before drinking it). Emotions, therefore, not only influence immediate actions but also serve as an important motivational basis for future behaviors.

Interpersonal Functions of Emotion

Emotions are expressed both verbally through words and nonverbally through facial expressions, voices, gestures, body postures, and movements. We are constantly expressing emotions when interacting with others, and others can reliably judge those emotional expressions (Elfenbein & Ambady, 2002; Matsumoto, 2001); thus, emotions have signal value to others and influence others and our social interactions. Emotions and their expressions communicate information to others about our feelings, intentions, relationship with the target of the emotions, and the environment. Because emotions have this communicative signal value, they help solve social problems by evoking responses from others, by signaling the nature of interpersonal relationships, and by providing incentives for desired social behavior (Keltner, 2003).

Emotions can act as signals to our friends and partners, conveying information about the quality of the relationship. [Image: mynameisharsha, https://goo.gl/HY2XgV, CC BY-SA 2.0, https://goo.gl/rxiUsF]

Emotional Expressions Facilitate Specific Behaviors in Perceivers

Because facial expressions of emotion are universal social signals, they contain meaning not only about the expressor's psychological state but also about that person's intent and subsequent behavior. This information affects what the perceiver is likely to do. People observing fearful faces, for instance, are more likely to produce approach-related behaviors, whereas people who observe angry faces are more likely to produce avoidance-related behaviors (Marsh, Ambady, & Kleck, 2005). Even subliminal presentation of smiles produces increases in how much beverage people pour and consume and how much they are willing to pay for it; presentation of angry faces decreases these behaviors (Winkielman, Berridge, & Wilbarger, 2005). Also, emotional displays evoke specific, complementary emotional responses from observers; for example, anger evokes fear in others (Dimberg & Ohman, 1996; Esteves, Dimberg, & Ohman, 1994), whereas distress evokes sympathy and aid (Eisenberg et al., 1989).

Emotional Expressions Signal the Nature of Interpersonal Relationships

Emotional expressions provide information about the nature of the relationships among interactants. Some of the most important and provocative set of findings in this area come from studies involving married couples (Gottman & Levenson, 1992; Gottman, Levenson, & Woodin, 2001). In this research, married couples visited a laboratory after having not seen each other for 24 hours, and then engaged in intimate conversations about daily events or issues of conflict. Discrete expressions of contempt, especially by the men, and disgust, especially by the women, predicted later marital dissatisfaction and even divorce.

Emotional Expressions Provide Incentives for Desired Social Behavior

Facial expressions of emotion are important regulators of social interaction. In the developmental literature, this concept has been investigated under the concept of social referencing (Klinnert, Campos, & Sorce, 1983); that is, the process whereby infants seek out information from others to clarify a situation and then use that information to act. To date, the strongest demonstration of social referencing comes from work on the visual cliff. In the first study to investigate this concept, Campos and colleagues (Sorce, Emde, Campos, & Klinnert, 1985) placed mothers on the far end of the "cliff" from the infant. Mothers first smiled to the infants and placed a toy on top the safety glass to attract them; infants invariably began crawling to their mothers. When the infants were in the center of the table, however, the mother then posed an expression of fear, sadness, anger, interest, or joy. The results were clearly different for the different faces; no infant crossed the table when the mother showed fear; only 6% did when the mother posed anger, 33% crossed when the mother posed sadness, and approximately 75% of the infants crossed when the mother posed joy or interest.

Other studies provide similar support for facial expressions as regulators of social interaction. In one study (Bradshaw, 1986), experimenters posed facial expressions of neutral, anger, or disgust toward babies as they moved toward an object and measured the amount of inhibition the babies showed in touching the object. The results for 10- and 15-month olds were the same: anger produced the greatest inhibition, followed by disgust, with neutral the least. This study was later replicated (Hertenstein & Campos, 2004) using joy and disgust expressions, altering the method so that the infants were not allowed to touch the toy (compared with a distractor object) until one hour after exposure to the expression. At 14 months of age, significantly more infants touched the toy when they saw joyful expressions, but fewer touched the toy when the infants saw disgust.

Social and Cultural Functions of Emotion

Although there are cultural differences in the display of emotion, almost all infants start showing emotion such as smiling or reacting to their caretaker as early as 6 weeks after their birth. [Image: vgm8383, https://goo.gl/jgfRDN, CC BY-NC 2.0, https://goo.gl/VnKlK8]

If you stop to think about many things we take for granted in our daily lives, we cannot help but come to the conclusion that modern human life is a colorful tapestry of many groups and individual lives woven together in a complex yet functional way. For example, when you're hungry, you might go to the local grocery store and buy some food. Ever stop to think about how you're able to do that? You might buy a banana that was grown in a field in southeast Asia being raised by farmers there, where they planted the tree, cared for it, and picked the fruit. They probably handed that fruit off to a distribution chain that allowed multiple people somewhere to use tools such as cranes, trucks, cargo bins, ships or airplanes (that were also created by multiple people somewhere) to bring that banana to your store. The store had people to care for that banana until you came and got it and to barter with you for it (with your money). You may have gotten to the store riding a vehicle that was produced somewhere else in the world by others, and you were probably wearing clothes produced by some other people somewhere else.

Thus, human social life is complex. Individuals are members of multiple groups, with multiple social roles, norms, and expectations, and people move rapidly in and out of the multiple groups of which they are members. Moreover, much of human social life is unique because it revolves around cities, where many people of disparate backgrounds come together. This creates the enormous potential for social chaos, which can easily occur if individuals are not coordinated well and relationships not organized systematically.

One of the important functions of culture is to provide this necessary coordination and organization. Doing so allows individuals and groups to negotiate the social complexity of human social life, thereby maintaining social order and preventing social chaos. Culture does this by providing a meaning and information system to its members, which is shared by a group and transmitted across generations, that allows the group to meet basic needs of survival, pursue happiness and well-being, and derive meaning from life (Matsumoto & Juang, 2013). Culture is what allowed the banana from southeast Asia to appear on your table.

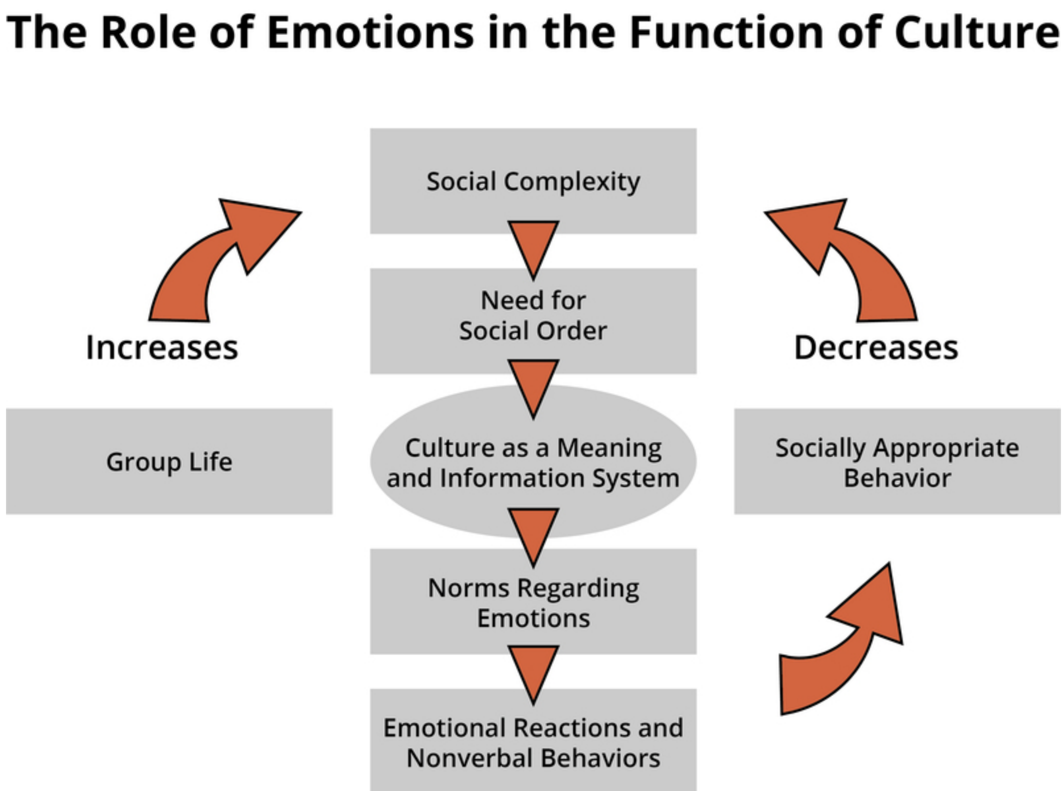


Figure 1: The Role of Emotions in the Function of Culture

Cultural transmission of the meaning and information system to its members is, therefore, a crucial aspect of culture. One of the ways this transmission occurs is through the development of worldviews (including attitudes, values, beliefs, and norms) related to emotions (Matsumoto & Hwang, 2013; Matsumoto et al., 2008). Worldviews related to emotions provide guidelines for desirable emotions that facilitate norms for regulating individual behaviors and interpersonal relationships. Our cultural backgrounds tell us which emotions are ideal to have, and which are not (Tsai, Knutson, & Fung, 2006). The cultural transmission of information related to emotions occurs in many ways, from childrearers to children, as well as from the cultural products available in our world, such as books, movies, ads, and the like (Schönpflug, 2009; Tsai, Louie, Chen, & Uchida, 2007).

Cultures also inform us about what to do with our emotions—that is, how to manage or modify them—when we experience them. One of the ways in which this is done is through the management of our emotional expressions through **cultural display rules** (Friesen, 1972). These are rules that are learned early in life that specify the management and modification of our emotional expressions according to social circumstances. Thus, we learn that "big boys don't cry" or to laugh at the boss's jokes even though they're not funny. By affecting how individuals express their emotions, culture also influences how people experience them as well.

Cultural display rules teach us how to manage our emotions. For example, in many Asian countries children are taught to mute their emotions, especially negative emotions like anger. [Image: john.gillespie, https://goo.gl/gTdPYb, CC BY-SA 3.0, https://goo.gl/eLCn2O]

Because one of the major functions of culture is to maintain social order in order to ensure group efficiency and thus survival, cultures create worldviews, rules, guidelines, and norms concerning emotions because emotions have important intra- and interpersonal functions, as described above, and are important motivators of behavior. Norms concerning emotion and its regulation in all cultures serve the purpose of maintaining social order. Cultural worldviews and norms help us manage and modify our emotional reactions (and thus behaviors) by helping us to have certain kinds of emotional experiences in the first place and by managing our reactions and subsequent behaviors once we have them. By doing so, our culturally moderated emotions can help us engage in socially appropriate behaviors, as defined by our cultures, and thus reduce social complexity and increase social order, avoiding social chaos. All of this allows us to live relatively harmonious and constructive lives in groups. If cultural worldviews and norms about emotions did not exist, people would just run amok having all kinds of emotional experiences, expressing their emotions and then behaving in all sorts of unpredictable and potentially harmful ways. If that were the case, it would be very difficult for groups and societies to function effectively, and even for humans to survive as a species, if emotions were not regulated in culturally defined ways for the common, social good. Thus, emotions play a critical role in the successful functioning of any society and culture.

Outside Resources

Alberta, G. M., Rieckmann, T. R., & Rush, J. D. (2000). Issues and recommendations for teaching an ethnic/culture-based course. Teaching of Psychology, 27,102-107. doi:10.1207/S15328023TOP2702_05
http://top.sagepub.com/content/27/2/102.short

CrashCourse (2014, August 4). Feeling all the feels: Crash course psychology #25. [Video file]. Retrieved from:
https://www.youtube.com/watch?v=gAMbkJk6gnE

Hughesm A. (2011). Exercises and demonstrations to promote student engagement in motivation and courses. In R. Miller, E. Balcetis, S. Burns, D. Daniel, B. Saville, & W. Woody (Eds.), Promoting Student Engagement: Volume 2: Activities, Exercises and Demonstrations for Psychology Courses. (pp. 79-82) Washington DC, Society for the Teaching of Psychology, American Psychological Association.
http://teachpsych.org/ebooks/pse2011/vol2/index.php

Johnston, E., & Olson, L. (2015). The feeling brain: The biology and psychology of emotions. New York, NY: W.W. Norton & Company.
http://books.wwnorton.com/books/The-Feeling-Brain/

NPR News: Science Of Sadness And Joy: 'Inside Out' Gets Childhood Emotions Right
http://www.npr.org/sections/health-shots/2015/06/13/413980258/science-of-sadness-and-joy-inside-out-gets-childhood-emotions-right

Online Psychology Laboratory: Motivation and Emotion resources
http://opl.apa.org/Resources.aspx#Motivation

Web: See how well you can read other people's facial expressions of emotion
http://www.humintell.com/free-demos/

Discussion Questions

1. When emotions occur, why do they simultaneously activate certain physiological and psychological systems in the body and deactivate others?

2. Why is it difficult for people to act rationally and think happy thoughts when they are angry? Conversely, why is it difficult to remember sad memories or have sad thoughts when people are happy?
3. You're walking down a deserted street when you come across a stranger who looks scared. What would you say? What would you do? Why?
4. You're walking down a deserted street when you come across a stranger who looks angry. What would you say? What would you do? Why?
5. Think about the messages children receive from their environment (such as from parents, mass media, the Internet, Hollywood movies, billboards, and storybooks). In what ways do these messages influence the kinds of emotions that children should and should not feel?

Vocabulary

Cultural display rules
These are rules that are learned early in life that specify the management and modification of emotional expressions according to social circumstances. Cultural display rules can work in a number of different ways. For example, they can require individuals to express emotions "as is" (i.e., as they feel them), to exaggerate their expressions to show more than what is actually felt, to tone down their expressions to show less than what is actually felt, to conceal their feelings by expressing something else, or to show nothing at all.

Interpersonal
This refers to the relationship or interaction between two or more individuals in a group. Thus, the interpersonal functions of emotion refer to the effects of one's emotion on others, or to the relationship between oneself and others.

Intrapersonal
This refers to what occurs within oneself. Thus, the intrapersonal functions of emotion refer to the effects of emotion to individuals that occur physically inside their bodies and psychologically inside their minds.

Social and cultural
Society refers to a system of relationships between individuals and groups of individuals; culture refers to the meaning and information afforded to that system that is transmitted across generations. Thus, the social and cultural functions of emotion refer to the effects that emotions have on the functioning and maintenance of societies and cultures.

Social referencing
This refers to the process whereby individuals look for information from others to clarify a situation, and then use that information to act. Thus, individuals will often use the emotional expressions of others as a source of information to make decisions about their own behavior.

References

Baumeister, R. F., Vohs, K. D., DeWall, N., & Zhang, L. (2007). How emotion shapes behavior: Feedback, anticipation, and reflection, rather than direct causation. *Personality and Social Psychology Review, 11*(2), 167–203.

Bradshaw, D. (1986). *Immediate and prolonged effectiveness of negative emotion expressions in inhibiting infants' actions* (Unpublished doctoral dissertation). Berkeley, CA: University of California, Berkeley.

Cannon, W. B. (1927). The James–Lange theory of emotions: A critical examination and an alternative theory. *American Journal of Psychology, 39*, 106–124.

Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 91–115). New York, NY: Guilford Press.

Darwin, C. (1872). *The expression of emotion in man and animals*. New York, NY: Oxford University Press.

Dimberg, U., & Ohman, A. (1996). Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation & Emotion, 20*(2), 149–182.

Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and distress to prosocial behavior: A multimethod study. *Journal of Personality and Social Psychology, 57*, 55–66.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 205–235.

Esteves, F., Dimberg, U., & Ohman, A. (1994). Automatically elicited fear: Conditioned skin conductance responses to masked facial expressions. *Cognition and Emotion, 8*(5), 393–413.

Friesen, W. V. (1972). *Cultural differences in facial expressions in a social situation: An experimental test of the concept of display rules* (Unpublished doctoral dissertation). San Francisco, CA: University of California, San Francisco.

Gottman, J. M., & Levenson, R. W. (1992). Marital processes predictive of later dissolution: Behavior, physiology, and health. *Journal of Personality and Social Psychology, 63*(2), 221–223.

Gottman, J. M., Levenson, R. W., & Woodin, E. (2001). Facial expressions during marital conflict. *Journal of Family Communication, 1*, 37–57.

Hertenstein, M. J., & Campos, J. J. (2004). The retention effects of an adult's emotional displays on infant behavior. *Child Development, 75*(2), 595–613.

James, W. (1890). *The principles of psychology*. New York, NY: Holt.

Keltner, D. (2003). Expression and the course of life: Studies of emotion, personality, and psychopathology from a social-functional perspective. In P. Ekman, J. Campos, R. J. Davidson, & F.B.M. De Waal (Eds.), *Emotions inside out: 130 years after Darwin's "The expression of the emotions in man and animals"* (Vol. 1000, pp. 222–243). New York, NY: New York Academy of Sciences.

Klinnert, M. D., Campos, J. J., & Sorce, J. F. (1983). Emotions as behavior regulators: Social referencing in infancy. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (pp. 57–86). New York, NY: Academic Press.

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition and Emotion, 13*(5), 481–504.

Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). *The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. Emotion, 5*(1), 119–124.

Matsumoto, D. (2001). Culture and emotion. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 171–194). New York, NY: Oxford University Press.

Matsumoto, D., & Hwang, H. C. (2013). Assessing cross-cultural competence: A review of available tests. *Journal of Cross-Cultural Psychology, 44*(6), 849–873.

Matsumoto, D., & Juang, L. (2013). *Culture and psychology* (5th ed.). Belmont, CA: Cengage.

Matsumoto, D., & Wilson, J. (2008). Culture, emotion, and motivation. In R. M. Sorrentino & S. Yamaguchi (Eds.), *Handbook of motivation and cognition across cultures* (pp. 541–563). New York, NY: Elsevier.

Matsumoto, D., Hirayama, S., & LeRoux, J. A. (2006). Psychological skills related to adjustment. In P.T.P. Wong & L.C.J. Wong (Eds.), *Handbook of multicultural perspectives on stress and coping* (pp. 387–405). New York, NY: Springer.

Matsumoto, D., Yoo, S. H., Nakagawa, S., Alexandre, J., Altarriba, J., Anguas-Wong, A. M., et al. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology, 94*(6), 925–937.

Schönpflug, U. (Ed.). (2009). *Cultural transmission: Developmental, psychological, social and methodological aspects*. New York, NY: Cambridge University Press.

Sorce, J. F., Emde, J. J., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology, 21*, 195–200.

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 114–137). New York, NY: The Guilford Press.

Tsai, J. L., Knutson, B., & Fung, H. H. (2006). Cultural variation in affect valuation. *Journal of*

Personality and Social Psychology, 90(2), 288–307.

Tsai, J. L., Louie, J. Y., Chen, E. E., & Uchida, Y. (2007). Learning what feelings to desire: Socialization of ideal affect through children's storybooks. *Personality and Social Psychology Bulletin, 33*(1), 17–30.

Wang, Q., & Ross, M. (2007). Culture and memory. In S. Kitayama & D. Cohen (Eds.), *Handbook of cultural psychology* (pp. 645–667). New York, NY: Guilford.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin, 31*(1), 121–135.

35
Drive States

Sudeep Bhatia & George Loewenstein

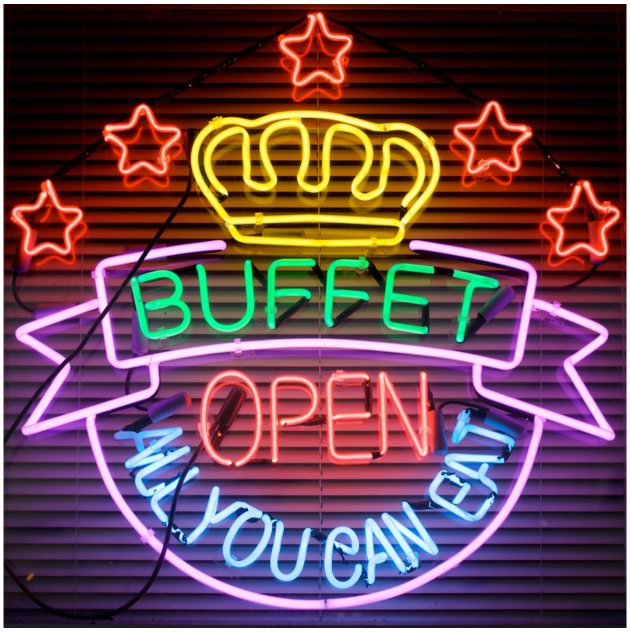
Our thoughts and behaviors are strongly influenced by affective experiences known as drive states. These drive states motivate us to fulfill goals that are beneficial to our survival and reproduction. This module provides an overview of key drive states, including information about their neurobiology and their psychological effects.

Learning Objectives

- Identify the key properties of drive states
- Describe biological goals accomplished by drive states
- Give examples of drive states
- Outline the neurobiological basis of drive states such as hunger and arousal
- Discuss the main moderators and determinants of drive states such as hunger and arousal

Introduction

What is the longest you've ever gone without eating? A couple of hours? An entire day? How did it feel? Humans rely critically on food for nutrition and energy, and the absence of food can create drastic changes, not only in physical appearance, but in thoughts and behaviors. If you've ever fasted for a day, you probably noticed how hunger can take over your mind, directing your attention to foods you could be eating (a cheesy slice of pizza, or perhaps some sweet, cold ice cream), and motivating you to obtain and consume these foods. And once you have eaten and your hunger has been satisfied, your thoughts and behaviors return to normal.



Hunger is among our most basic motivators. [Image: Jeremy Brooks, https://goo.gl/XrFG2W, CC BY-NC 2.0, https://goo.gl/VnKlK8]

Hunger is a **drive state**, an affective experience (something you feel, like the sensation of being tired or hungry) that motivates organisms to fulfill goals that are generally beneficial to their survival and reproduction. Like other drive states, such as thirst or sexual arousal, hunger has a profound impact on the functioning of the mind. It affects psychological processes, such as perception, attention, emotion, and motivation, and influences the behaviors that these processes generate.

Key Properties of Drive States

Drive states differ from other affective or emotional states in terms of the biological functions they accomplish. Whereas all affective states possess valence (i.e., they are positive or negative) and serve to motivate approach or avoidance behaviors (Zajonc, 1998), drive states are unique in that they generate behaviors that result in specific benefits for the body. For example, hunger directs individuals to eat foods that increase blood sugar levels in the body, while thirst causes individuals to drink fluids that increase water levels in the body.

Different drive states have different triggers. Most drive states respond to both internal and external cues, but the combinations of internal and external cues, and the specific types of cues, differ between drives. Hunger, for example, depends on internal, visceral signals as well as sensory signals, such as the sight or smell of tasty food. Different drive states also result in different cognitive and emotional states, and are associated with different behaviors. Yet despite these differences, there are a number of properties common to all drive states.

Homeostasis

Humans, like all organisms, need to maintain a stable state in their various physiological systems. For example, the excessive loss of body water results in dehydration, a dangerous and potentially fatal state. However, too much water can be damaging as well. Thus, a moderate and stable level of body fluid is ideal. The tendency of an organism to maintain this stability across all the different physiological systems in the body is called **homeostasis**.

Homeostasis is maintained via two key factors. First, the state of the system being regulated must be monitored and compared to an ideal level, or a **set point**. Second, there need to be mechanisms for moving the system back to this set point—that is, to restore homeostasis when deviations from it are detected. To better understand this, think of the thermostat in your own home. It detects when the current temperature in the house is different than the temperature you have it set at (i.e., the set point). Once the thermostat recognizes the difference, the heating or air conditioning turns on to bring the overall temperature back to the designated level.

Many homeostatic mechanisms, such as blood circulation and immune responses, are automatic and nonconscious. Others, however, involve deliberate action. Most drive states motivate action to restore homeostasis using both "punishments" and "rewards." Imagine that these homeostatic mechanisms are like molecular parents. When you behave poorly by departing from the set point (such as not eating or being somewhere too cold), they raise their voice at you. You experience this as the bad feelings, or "punishments," of hunger, thirst, or feeling too cold or too hot. However, when you behave well (such as eating nutritious foods when hungry), these homeostatic parents reward you with the pleasure that comes from any activity that moves the system back toward the set point. For example, when body temperature declines below the set point, any activity that helps to restore homeostasis (such as putting one's hand in warm water) feels pleasurable; and likewise, when body temperature rises above the set point, anything that cools it feels pleasurable.

The body needs homeostasis and motivates us - through both pleasure and pain - to stay in balance. [Image: ashleigh290, https://goo.gl/yXQtEC, CC-BY 2.0, https://goo.gl/BRvSA7]

The Narrowing of Attention

As drive states intensify, they direct attention toward elements, activities, and forms of consumption that satisfy the biological needs associated with the drive. Hunger, for example, draws attention toward food. Outcomes and objects that are not related to satisfying hunger lose their value (Easterbrook, 1959). For instance, has anyone ever invited you to do a fun activity while you were hungry? Likely your response was something like: “I’m not doing anything until I eat first.” Indeed, at a sufficient level of intensity, individuals will sacrifice almost any quantity of goods that do not address the needs signaled by the drive state. For example, cocaine addicts, according to Gawin (1991:1581), “report that virtually all thoughts are focused on cocaine during binges; nourishment, sleep, money, loved ones, responsibility, and survival lose all significance.”

Drive states also produce a second form of attention-narrowing: a collapsing of time-perspective toward the present. That is, they make us impatient. While this form of attention-narrowing is particularly pronounced for the outcomes and behaviors directly related to the biological function being served by the drive state at issue (e.g., “I need food *now*”), it applies to general concerns for the future as well. Ariely and Loewenstein (2006), for example, investigated the impact of sexual arousal on the thoughts and behaviors of a sample of male undergraduates. These undergraduates were lent laptop computers that they took to their private residences, where they answered a series of questions, both in normal states and in states of high sexual arousal. Ariely and Loewenstein found that being sexually aroused made people extremely impatient for both sexual outcomes and for outcomes in other domains, such as those involving money. In another study Giordano et al. (2002) found that heroin addicts were more impatient with respect to heroin when they were craving it than when they were not. More surprisingly, they were also more impatient toward money (they valued delayed money less) when they were actively craving heroin.

Yet a third form of attention-narrowing involves thoughts and outcomes related to the self versus others. Intense drive states tend to narrow one’s focus inwardly and to undermine altruism—or the desire to do good for others. People who are hungry, in pain, or craving drugs tend to be selfish. Indeed, popular interrogation methods involve depriving individuals of sleep, food, or water, so as to trigger intense drive states leading the subject of the interrogation to divulge information that may betray comrades, friends, and family (Biderman, 1960).

Two Illustrative Drive States

Thus far we have considered drive states abstractly. We have discussed the ways in which they relate to other affective and motivational mechanisms, as well as their main biological purpose and general effects on thought and behavior. Yet, despite serving the same broader goals, different drive states are often remarkably different in terms of their specific properties. To understand some of these specific properties, we will explore two different drive states

Current Controversy

In 2005, the American Psychological Association (APA) issued a report concluding that psychologists could ethically play a role in the interrogation of people captured in Afghanistan and elsewhere. In 2014, following critical media publicity documenting the APA's involvement in torture, the APA commissioned a law firm to independently investigate APA involvement in interrogation. The firm's report was damaging to the APA because it suggested that APA leaders colluded with the Department of Defense, CIA, and other government officials not only to aid in interrogation itself, but to provide justification for government guidelines that defined torture (which is banned by international treaties signed by the U.S.) in a narrow fashion that excluded, for example so-called "stress positions" and sleep deprivation.

Critical Questions

1. ***Do you think that manipulating drive states, such as the need for sleep, constitutes torture?***
2. ***How do you think research on drive states should inform the definition of "torture" and our definition of ethical interrogation techniques?***

See the full Hoffman Report here - http://www.apa.org/independent-review/APA-FINAL-Report-7.2.15.pdf

For more coverage - http://www.nytimes.com/2015/07/11/us/psychologists-shielded-us-torture-program-report-finds.html

that play very important roles in determining behavior, and in ensuring human survival: hunger and sexual arousal.

Hunger

Hunger is a classic example of a drive state, one that results in thoughts and behaviors related to the consumption of food. Hunger is generally triggered by low glucose levels in the blood (Rolls, 2000), and behaviors resulting from hunger aim to restore homeostasis regarding those glucose levels. Various other internal and external cues can also cause hunger. For example, when fats are broken down in the body for energy, this initiates a chemical cue that the body should search for food (Greenberg, Smith, & Gibbs, 1990). External cues include the time of day, estimated time until the next feeding (hunger increases immediately prior to food consumption), and the sight, smell, taste, and even touch of food and food-related stimuli. Note that while hunger is a generic feeling, it has nuances that can provoke the eating of specific foods that correct for nutritional imbalances we may not even be conscious of. For

External cues, like the sight and smell of food, can ignite feelings of hunger. [Image: Marco Verch, https://goo.gl/c4TC5A, CC BY 2.0, https://goo.gl/BRvSA7]

example, a couple who was lost adrift at sea found they inexplicably began to crave the eyes of fish. Only later, after they had been rescued, did they learn that fish eyes are rich in vitamin C—a very important nutrient that they had been depleted of while lost in the ocean (Walker, 2014).

The **hypothalamus** (located in the lower, central part of the brain) plays a very important role in eating behavior. It is responsible for synthesizing and secreting various hormones. The lateral hypothalamus (LH) is concerned largely with hunger and, in fact, lesions (i.e., damage) of the LH can eliminate the desire for eating entirely—to the point that animals starve themselves to death unless kept alive by force feeding (Anand & Brobeck, 1951). Additionally, artificially stimulating the LH, using electrical currents, can generate eating behavior if food is available (Andersson, 1951).

Activation of the LH can not only increase the desirability of food but can also reduce the desirability of nonfood-related items. For example, Brendl, Markman, and Messner (2003) found that participants who were given a handful of popcorn to trigger hunger not only had higher ratings of food products, but also had lower ratings of nonfood products—compared with participants whose appetites were not similarly primed. That is, because eating had become more important, other non-food products lost some of their value.

Hunger is only part of the story of when and why we eat. A related process, satiation, refers to the decline of hunger and the eventual termination of eating behavior. Whereas the feeling of hunger gets you to start eating, the feeling of satiation gets you to stop. Perhaps surprisingly, hunger and satiation are two distinct processes, controlled by different circuits in the brain and triggered by different cues. Distinct from the LH, which plays an important role in hunger, the ventromedial hypothalamus (VMH) plays an important role in satiety. Though lesions of the VMH can cause an animal to overeat to the point of obesity, the relationship between the LH and the VMB is quite complicated. Rats with VMH lesions can also be quite finicky about their food (Teitelbaum, 1955).

Other brain areas, besides the LH and VMH, also play important roles in eating behavior. The sensory cortices (visual, olfactory, and taste), for example, are important in identifying food items. These areas provide informational value, however, not hedonic evaluations. That is, these areas help tell a person what is good or safe to eat, but they don't provide the pleasure (or hedonic) sensations that *actually* eating the food produces. While many sensory functions are roughly stable across different psychological states, other functions, such as the detection of food-related stimuli, are enhanced when the organism is in a hungry drive state.

After identifying a food item, the brain also needs to determine its **reward value**, which affects the organism's motivation to consume the food. The reward value ascribed to a particular item is, not surprisingly, sensitive to the level of hunger experienced by the organism. The hungrier you are, the greater the reward value of the food. Neurons in the areas where reward values are processed, such as the orbitofrontal cortex, fire more rapidly at the sight or taste of food when the organism is hungry relative to if it is satiated.

Sexual Arousal

A second drive state, especially critical to reproduction, is sexual arousal. Sexual arousal results in thoughts and behaviors related to sexual activity. As with hunger, it is generated by a large range of internal and external mechanisms that are triggered either after the extended absence of sexual activity or by the immediate presence and possibility of sexual activity (or by cues commonly associated with such possibilities). Unlike hunger, however, these mechanisms can differ substantially between males and females, indicating important evolutionary differences in the biological functions that sexual arousal serves for different sexes.

Unlike other drive states the mechanisms that trigger sexual arousal are not the same for men and women. [Image: Matthew Romack, https://goo.gl/IUbbk0, CC BY 2.0, https://goo.gl/BRvSA7]

Sexual arousal and pleasure in males, for example, is strongly related to the **preoptic area**, a region in the anterior hypothalamus (or the front of the hypothalamus). If the preoptic area is damaged, male sexual behavior is severely impaired. For example, rats that have had prior sexual experiences will still seek out sexual partners after their preoptic area is lesioned.

However, once having secured a sexual partner, rats with lesioned preoptic areas will show no further inclination to actually initiate sex.

For females, though, the preoptic area fulfills different roles, such as functions involved with eating behaviors. Instead, there is a different region of the brain, the ventromedial hypothalamus (the lower, central part) that plays a similar role for females as the preoptic area does for males. Neurons in the ventromedial hypothalamus determine the excretion of estradiol, an estrogen hormone that regulates sexual receptivity (or the willingness to accept a sexual partner). In many mammals, these neurons send impulses to the periaqueductal gray (a region in the midbrain) which is responsible for defensive behaviors, such as freezing immobility, running, increases in blood pressure, and other motor responses. Typically, these defensive responses might keep the female rat from interacting with the male one. However, during sexual arousal, these defensive responses are weakened and lordosis behavior, a physical sexual posture that serves as an invitation to mate, is initiated (Kow and Pfaff, 1998). Thus, while the preoptic area encourages males to engage in sexual activity, the ventromedial hypothalamus fulfills that role for females.

Other differences between males and females involve overlapping functions of neural modules. These neural modules often provide clues about the biological roles played by sexual arousal and sexual activity in males and females. Areas of the brain that are important for male sexuality overlap to a great extent with areas that are also associated with aggression. In contrast, areas important for female sexuality overlap extensively with those that are also connected to nurturance (Panksepp, 2004).

One region of the brain that seems to play an important role in sexual pleasure for both males and females is the septal nucleus, an area that receives reciprocal connections from many other brain regions, including the hypothalamus and the amygdala (a region of the brain primarily involved with emotions). This region shows considerable activity, in terms of rhythmic spiking, during sexual orgasm. It is also one of the brain regions that rats will most reliably voluntarily self-stimulate (Olds & Milner, 1954). In humans, placing a small amount of acetylcholine into this region, or stimulating it electrically, has been reported to produce a feeling of imminent orgasm (Heath, 1964).

Conclusion

Drive states are evolved motivational mechanisms designed to ensure that organisms take self-beneficial actions. In this module, we have reviewed key properties of drive states, such as homeostasis and the narrowing of attention. We have also discussed, in some detail, two important drive states—hunger and sexual arousal—and explored their underlying neurobiology and the ways in which various environmental and biological factors affect their properties.

There are many drive states besides hunger and sexual arousal that affect humans on a daily basis. Fear, thirst, exhaustion, exploratory and maternal drives, and drug cravings are all drive states that have been studied by researchers (see e.g., Buck, 1999; Van Boven & Loewenstein, 2003). Although these drive states share some of the properties discussed in this module, each also has unique features that allow it to effectively fulfill its evolutionary function.

One key difference between drive states is the extent to which they are triggered by internal as opposed to external stimuli. Thirst, for example, is induced both by decreased fluid levels and an increased concentration of salt in the body. Fear, on the other hand, is induced by perceived threats in the external environment. Drug cravings are triggered both by internal homeostatic mechanisms and by external visual, olfactory, and contextual cues. Other drive states, such as those pertaining to maternity, are triggered by specific events in the organism's life. Differences such as these make the study of drive states a scientifically interesting and important endeavor. Drive states are rich in their diversity, and many questions involving their neurocognitive underpinnings, environmental determinants, and behavioral effects, have yet to be answered.

One final thing to consider, not discussed in this module, relates to the real-world consequences of drive states. Hunger, sexual arousal, and other drive states are all psychological mechanisms that have evolved gradually over millions of years. We share these drive states not only with our human ancestors but with other animals, such as monkeys, dogs, and rats. It is not surprising then that these drive states, at times, lead us to behave in ways that are ill-suited to our modern lives. Consider, for example, the obesity epidemic that is affecting countries around the world. Like other diseases of affluence, obesity is a product of drive states that are too easily fulfilled: homeostatic mechanisms that once worked well when food was scarce now backfire when meals rich in fat and sugar are readily available. Unrestricted sexual arousal can have similarly perverse effects on our well-being. Countless politicians have sacrificed their entire life's work (not to mention their marriages) by indulging adulterous sexual impulses toward colleagues, staffers, prostitutes, and others over whom they have social or financial power. It not an overstatement to say that many problems of the 21st century, from school massacres to obesity to drug addiction, are influenced by the mismatch between our drive states and our uniquely modern ability to fulfill them at a moment's notice.

Outside Resources

Web: An open textbook chapter on homeostasis
http://en.wikibooks.org/wiki/Human_Physiology/Homeostasis

Web: Motivation and emotion in psychology
http://allpsych.com/psychology101/motivation_emotion.html

Web: The science of sexual arousal
http://www.apa.org/monitor/apr03/arousal.aspx

Discussion Questions

1. The ability to maintain homeostasis is important for an organism's survival. What are the ways in which homeostasis ensures survival? Do different drive states accomplish homeostatic goals differently?
2. Drive states result in the narrowing of attention toward the present and toward the self. Which drive states lead to the most pronounced narrowing of attention toward the present? Which drive states lead to the most pronounced narrowing of attention toward the self?
3. What are important differences between hunger and sexual arousal, and in what ways do these differences reflect the biological needs that hunger and sexual arousal have been evolved to address?
4. Some of the properties of sexual arousal vary across males and females. What other drives states affect males and females differently? Are there drive states that vary with other differences in humans (e.g., age)?

Vocabulary

Drive state
Affective experiences that motivate organisms to fulfill goals that are generally beneficial to their survival and reproduction.

Homeostasis
The tendency of an organism to maintain a stable state across all the different physiological systems in the body.

Homeostatic set point
An ideal level that the system being regulated must be monitored and compared to.

Hypothalamus
A portion of the brain involved in a variety of functions, including the secretion of various hormones and the regulation of hunger and sexual arousal.

Lordosis
A physical sexual posture in females that serves as an invitation to mate.

Preoptic area
A region in the anterior hypothalamus involved in generating and regulating male sexual behavior.

Reward value
A neuropsychological measure of an outcome’s affective importance to an organism.

Satiation
The state of being full to satisfaction and no longer desiring to take on more.

References

Anand, B. K., & Brobeck, J. R. (1951). Hypothalamic control of food intake in rats and cats. *The Yale journal of biology and medicine, 24(2)*, 123.

Andersson, B. (1951). The effect and localization of electrical stimulation of certain parts of the brain stem in sheep and goats. *Acta Physiologica Scandinavica, 23*(1), 8–23.

Ariely, D., & Loewenstein, G. (2006). The heat of the moment: The effect of sexual arousal on sexual decision making. *Journal of Behavioral Decision Making, 19*(2), 87–98

Biderman, A. D. (1960). Social-psychological needs and “involuntary” behavior as illustrated by compliance in interrogation. *Sociometry, 23*(2), 120–147.

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.

Buck, R. (1999). The biological affects: A typology. *Psychological Review, 106*, 301–336.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*(3).

Gawin, F. H. (1991). Cocaine addiction: psychology and neurophysiology. *Science, 251*(5001), 1580–1586.

Giordano, L. A., Bickel, W. K., Loewenstein, G., Jacobs, E. A., Marsch, L., & Badger, G. J. (2002). Mild opioid deprivation increases the degree that opioid-dependent outpatients discount delayed heroin and money. *Psychopharmacology, 163*(2), 174–182.

Greenberg, D., Smith, G. P., & Gibbs, J. (1990). Intraduodenal infusions of fats elicit satiety in sham-feeding rats. *American Journal of Physiology-Regulatory, Integrative, and Comparative Physiology, 259*(1), 110–118.

Heath, R. G. (1964). Pleasure response of human subjects to direct stimulation of the brain: Physiologic and psychodynamic considerations. In R. G. Heath (Ed.), *The role of pleasure in behavior* (pp. 219–243). New York, NY: Hoeber.

Kow, L. M., & Pfaff, D. W. (1998). Mapping of neural and signal transduction pathways for lordosis in the search for estrogen actions on the central nervous system. *Behavioural Brain Research, 92*(2), 169–180

Olds, J., & Milner, P. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology, 47* (6), 419.

Panksepp, J. (2004). *Affective neuroscience: The foundations of human and animal emotions* (Vol. 4). New York, NY: Oxford University Press U.S.

Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10*(3), 284-294.

Teitelbaum, P. (1955). Sensory control of hypothalamic hyperphagia. *Journal of Comparative and Physiological Psychology, 48*(3), 156.

Van Boven, L., & Loewenstein, G. (2003). Projection of transient drive states. *Personality and Social Psychology Bulletin, 29*, 1159–1168.

Walker, P. (2014). Castaway's sea savvy could have helped him survive year adrift, says expert. The Guardian. Retried from http://www.theguardian.com/world/2014/feb/04/castaway-jose-salvador-alvarenga-survival-expert

Zajonc, R. B. (1998). Emotions. In Gilbert, D. T., Fiske, S. T., & Lindzey, G. (Eds.), *The handbook of social psychology* (Vols. 1 and 2, 4th ed., pp. 591–632). New York, NY: McGraw-Hill.

Psychological Disorders

36
History of Mental Illness

Ingrid G. Farreras

This module is divided into three parts. The first is a brief introduction to various criteria we use to define or distinguish between normality and abnormality. The second, largest part is a history of mental illness from the Stone Age to the 20th century, with a special emphasis on the recurrence of three causal explanations for mental illness; supernatural, somatogenic, and psychogenic factors. This part briefly touches upon trephination, the Greek theory of hysteria within the context of the four bodily humors, witch hunts, asylums, moral treatment, mesmerism, catharsis, the mental hygiene movement, deinstitutionalization, community mental health services, and managed care. The third part concludes with a brief description of the issue of diagnosis.

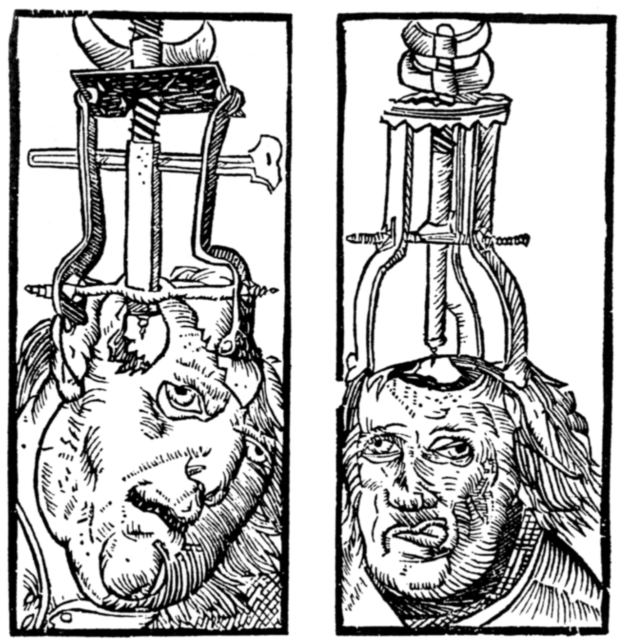
Learning Objectives

- Identify what the criteria used to distinguish normality from abnormality are.
- Understand the difference among the three main etiological theories of mental illness.
- Describe specific beliefs or events in history that exemplify each of these etiological theories (e.g., hysteria, humorism, witch hunts, asylums, moral treatments).
- Explain the differences in treatment facilities for the mentally ill (e.g., mental hospitals, asylums, community mental health centers).
- Describe the features of the "moral treatment" approach used by Chiarughi, Pinel, and Tuke.
- Describe the reform efforts of Dix and Beers and the outcomes of their work.
- Describe Kräpelin's classification of mental illness and the current DSM system.

History of Mental Illness

References to mental illness can be found throughout history. The evolution of mental illness, however, has not been linear or progressive but rather cyclical. Whether a behavior is considered normal or abnormal depends on the context surrounding the behavior and thus changes as a function of a particular time and culture. In the past, uncommon behavior or behavior that deviated from the sociocultural norms and expectations of a specific culture and period has been used as a way to silence or control certain individuals or groups. As a result, a less **cultural relativist** view of abnormal behavior has focused instead on whether behavior poses a threat to oneself or others or causes so much pain and suffering that it interferes with one's work responsibilities or with one's relationships with family and friends.

Throughout history there have been three general theories of the **etiology** of mental illness: **supernatural**, **somatogenic**, and **psychogenic**. Supernatural theories attribute mental illness to possession by evil or demonic spirits, displeasure of gods, eclipses, planetary gravitation, curses, and sin. Somatogenic theories identify disturbances in physical functioning resulting from either illness, genetic inheritance, or brain damage or imbalance. Psychogenic theories focus on traumatic or stressful experiences, **maladaptive** learned associations and cognitions, or distorted perceptions. Etiological theories of mental illness determine the care and treatment mentally ill individuals receive. As we will see below, an individual believed to be possessed by the devil will be viewed and treated differently from an individual believed to be suffering from an excess of yellow bile. Their treatments will also differ, from exorcism to blood-letting. The theories, however, remain the same. They coexist as well as recycle over time.



Engravings from 1525 showing trephination. It was believed that drilling holes in the skull could cure mental disorders. [Image: Peter Treveris, CC0 Public Domain, https://goo.gl/m25gce]

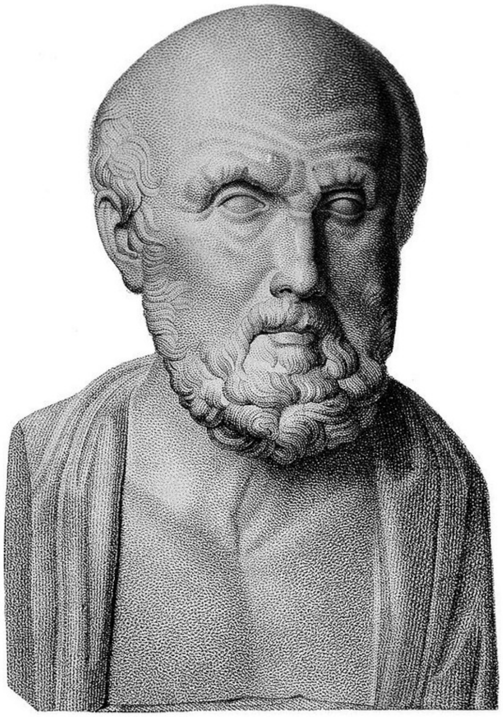
Trephination is an example of the earliest supernatural explanation for mental illness. Examination of prehistoric skulls and cave art from as early as 6500 BC has identified surgical drilling of holes in skulls to treat head injuries and epilepsy as well as to allow evil spirits trapped within the skull to be released (Restak, 2000). Around 2700 BC, Chinese medicine's concept of complementary positive and negative bodily forces ("yin and yang") attributed mental (and physical) illness to an imbalance between these forces. As such, a harmonious life that allowed for the proper balance of yin and yang and movement of vital air was essential (Tseng, 1973).

Mesopotamian and Egyptian papyri from 1900 BC describe women suffering from mental illness resulting from a wandering uterus (later named **hysteria** by the Greeks): The uterus could become dislodged and attached to parts of the body like the liver or chest cavity, preventing their proper functioning or producing varied and sometimes painful symptoms. As a result, the Egyptians, and later the Greeks, also employed a somatogenic treatment of strong smelling substances to guide the uterus back to its proper location (pleasant odors to lure and unpleasant ones to dispel).

Throughout classical antiquity we see a return to supernatural theories of demonic possession or godly displeasure to account for abnormal behavior that was beyond the person's control. Temple attendance with religious healing ceremonies and incantations to the gods were employed to assist in the healing process. Hebrews saw madness as punishment from God, so treatment consisted of confessing sins and repenting. Physicians were also believed to be able to comfort and cure madness, however.

Greek physicians rejected supernatural explanations of mental disorders. It was around 400 BC that Hippocrates (460–370 BC) attempted to separate superstition and religion from medicine by systematizing the belief that a deficiency in or especially an excess of one of the four essential bodily fluids (i.e., humors)—blood, yellow bile, black bile, and phlegm—was responsible for physical and mental illness. For example, someone who was too temperamental suffered from too much blood and thus blood-letting would be the necessary treatment. Hippocrates classified mental illness into one of four categories—epilepsy, mania, melancholia, and brain fever—and like other prominent physicians and philosophers of his time, he did not believe mental illness was shameful or that mentally ill individuals should be held accountable for their behavior. Mentally ill individuals were cared for at home by family members and the state shared no responsibility for their care. **Humorism** remained a recurrent somatogenic theory up until the 19th century.

While Greek physician Galen (AD 130–201) rejected the notion of a uterus having an **animistic soul**, he agreed with the notion that an imbalance of the four bodily fluids could cause mental illness. He also opened the door for psychogenic explanations for mental illness, however, by allowing for the experience of psychological stress as a potential cause of abnormality. Galen's psychogenic theories were ignored for centuries, however, as physicians attributed mental



Many of Hippocrates' medical theories are no longer practiced today. However, he pioneered medicine as an empirical practice and came up with the "Hippocratic oath," which all doctors must swear to before joining the profession (i.e., the promise to never intentionally harm a patient). [Image: Wellcome Images, https://goo.gl/dX21yj, CC BY 4.0, https://goo.gl/FJIuOM]

illness to physical causes throughout most of the millennium.

By the late Middle Ages, economic and political turmoil threatened the power of the Roman Catholic church. Between the 11th and 15th centuries, supernatural theories of mental disorders again dominated Europe, fueled by natural disasters like plagues and famines that lay people interpreted as brought about by the devil. Superstition, astrology, and alchemy took hold, and common treatments included prayer rites, relic touching, confessions, and atonement. Beginning in the 13th century the mentally ill, especially women, began to be persecuted as witches who were possessed. At the height of the witch hunts during the 15th through 17th centuries, with the Protestant Reformation having plunged Europe into religious strife, two Dominican monks wrote the *Malleus Maleficarum* (1486) as the ultimate manual to guide witch hunts. Johann Weyer and Reginald Scot tried to convince people in the mid- to late-16th century that accused witches were actually women with mental illnesses and that mental illness was not due to demonic possession but to faulty metabolism and disease, but the Church's Inquisition banned both of their writings. Witch-hunting did not decline until the 17th and 18th centuries, after more than 100,000 presumed witches had been burned at the stake (Schoeneman, 1977; Zilboorg & Henry, 1941).

Modern treatments of mental illness are most associated with the establishment of hospitals and **asylums** beginning in the 16th century. Such institutions' mission was to house and confine the mentally ill, the poor, the homeless, the unemployed, and the criminal. War and economic depression produced vast numbers of undesirables and these were separated from society and sent to these institutions. Two of the most well-known institutions, St. Mary of Bethlehem in London, known as Bedlam, and the Hôpital Général of Paris—which included La Salpêtrière, La Pitié, and La Bicêtre—began housing mentally ill patients in the mid-16th and 17th centuries. As confinement laws focused on protecting the public *from* the mentally ill, governments became responsible for housing and feeding undesirables in exchange for their personal liberty. Most inmates were institutionalized against their will, lived in filth and chained to walls, and were commonly exhibited to the public for a fee. Mental illness was nonetheless viewed somatogenically, so treatments were similar to those for physical illnesses: purges, bleedings, and emetics.

While inhumane by today's standards, the view of insanity at the time likened the mentally ill to animals (i.e., animalism) who did not have the capacity to reason, could not control themselves, were capable of violence without provocation, did not have the same physical sensitivity to pain or temperature, and could live in miserable conditions without complaint. As such, instilling fear was believed to be the best way to restore a disordered mind to reason.

By the 18th century, protests rose over the conditions under which the mentally ill lived, and the 18th and 19th centuries saw the growth of a more humanitarian view of mental illness. In 1785 Italian physician Vincenzo Chiarughi (1759–1820) removed the chains of patients at his St. Boniface hospital in Florence, Italy, and encouraged good hygiene and recreational and occupational training. More well known, French physician Philippe Pinel (1745–1826) and former patient Jean-Baptise Pussin created a "**traitement moral**" at La Bicêtre and the Salpêtrière in 1793 and 1795 that also included unshackling patients, moving them to well-aired, well-lit rooms, and encouraging purposeful activity and freedom to move about the grounds (Micale, 1985).

In England, humanitarian reforms rose from religious concerns. William Tuke (1732–1822) urged the Yorkshire Society of (Quaker) Friends to establish the York Retreat in 1796, where patients were guests, not prisoners, and where the standard of care depended on dignity and courtesy as well as the therapeutic and moral value of physical work (Bell, 1980).

Dorothea Dix worked to change the negative perceptions of people with mental illness and helped create institutions where they could receive compassionate care. [Image: State Archives of North Carolina, https://goo.gl/wRgGsi, no known copyright restrictions]

While America had asylums for the mentally ill—such as the Pennsylvania Hospital in Philadelphia and the Williamsburg Hospital, established in 1756 and 1773—the somatogenic theory of mental illness of the time—promoted especially by the father of America psychiatry, Benjamin Rush (1745–1813)—had led to treatments such as blood-letting, gyrators, and tranquilizer chairs. When Tuke's York Retreat became the model for half of the new private asylums established in the United States, however, psychogenic treatments such as compassionate care and physical labor became the hallmarks of the new American asylums, such as the Friends Asylum in Frankford, Pennsylvania, and the Bloomingdale Asylum in New York City, established in 1817 and 1821 (Grob, 1994).

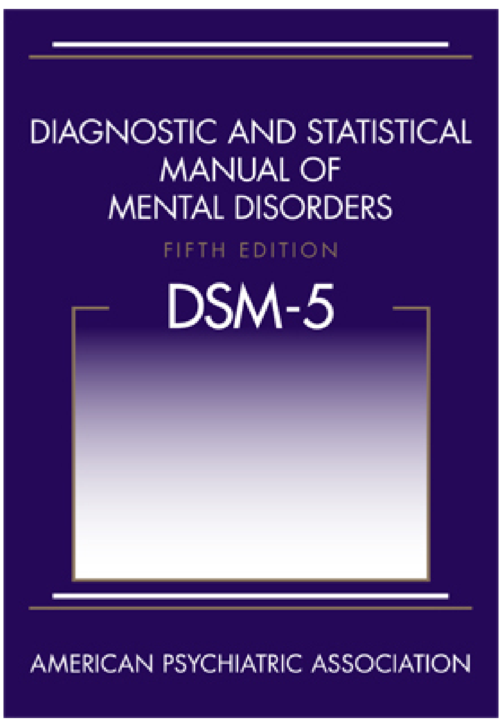
Moral treatment had to be abandoned in America in the second half of the 19th century, however, when these asylums became overcrowded and custodial in nature and could no longer provide the space nor attention necessary. When retired school teacher Dorothea Dix discovered the negligence that resulted from such conditions, she advocated for the establishment of state hospitals. Between 1840 and1880, she helped establish over 30 mental institutions in the United States and Canada (Viney & Zorich, 1982). By the late 19th century, moral treatment had given way to the mental hygiene movement, founded by former patient Clifford Beers with the publication of his 1908 memoir *A Mind That Found Itself*. Riding on Pasteur's breakthrough germ theory of the 1860s and 1870s and especially on the early 20th century discoveries of vaccines for cholera, syphilis, and typhus, the mental hygiene movement reverted to a somatogenic theory of mental illness.

European psychiatry in the late 18th century and throughout the 19th century, however, struggled between somatogenic and psychogenic explanations of mental illness, particularly hysteria, which caused physical symptoms such as blindness or paralysis with no apparent physiological explanation. Franz Anton Mesmer (1734–1815), influenced by contemporary discoveries in electricity, attributed hysterical symptoms to imbalances in a universal magnetic fluid found in individuals, rather than to a wandering uterus (Forrest, 1999). James Braid (1795–1860) shifted this belief in **mesmerism** to one in hypnosis, thereby proposing a psychogenic treatment for the removal of symptoms. At the time, famed Salpetriere Hospital neurologist Jean-Martin Charcot (1825–1893), and Ambroise Auguste Liébault (1823–1904) and Hyppolyte Bernheim (1840–1919) of the Nancy School in France, were engaged in a bitter etiological battle over hysteria, with Charcot maintaining that the hypnotic suggestibility underlying hysteria was a neurological condition while Liébault and Bernheim believed it to be a general trait that varied in the population. Josef Breuer (1842–1925) and Sigmund Freud (1856–1939) would resolve this dispute in favor of a psychogenic explanation for mental illness by treating hysteria through hypnosis, which eventually led to the **cathartic method** that became the precursor for psychoanalysis during the first half of the 20th century.

Psychoanalysis was the dominant psychogenic treatment for mental illness during the first half of the 20th century, providing the launching pad for the more than 400 different schools of psychotherapy found today (Magnavita, 2006). Most of these schools cluster around broader behavioral, cognitive, cognitive-behavioral, psychodynamic, and client-centered approaches to psychotherapy applied in individual, marital, family, or group formats. Negligible differences have been found among all these approaches, however; their efficacy in treating mental illness is due to factors shared among all of the approaches (not particular elements specific to each approach): the therapist-patient alliance, the therapist's allegiance to the therapy, therapist competence, and placebo effects (Luborsky et al., 2002; Messer & Wampold, 2002).

In contrast, the leading somatogenic treatment for mental illness can be found in the establishment of the first psychotropic medications in the mid-20th century. Restraints, electro-convulsive shock therapy, and lobotomies continued to be employed in American state institutions until the 1970s, but they quickly made way for a burgeoning pharmaceutical industry that has viewed and treated mental illness as a chemical imbalance in the brain.

Both etiological theories coexist today in what the psychological discipline holds as the biopsychosocial model of explaining human behavior. While individuals may be born with a genetic predisposition for a certain psychological disorder, certain psychological stressors need to be present for them to develop the disorder. Sociocultural factors such as sociopolitical or economic unrest, poor living conditions, or problematic interpersonal relationships are also viewed as contributing factors. However much we want to believe that we are above the treatments described above, or that the present is always the most enlightened time, let us not forget that our thinking today continues to reflect the same underlying somatogenic and psychogenic theories of mental illness discussed throughout this cursory 9,000-year history.



Up until the 1970's, homosexuality was included in the DSM as a psychological disorder. Thankfully, society and clinical understanding changed to recognize it didn't belong. [Image: Rene Walter, https://goo.gl/CcJAA1, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF]

Diagnosis of Mental Illness

Progress in the treatment of mental illness necessarily implies improvements in the diagnosis of mental illness. A standardized diagnostic classification system with agreed-upon definitions of psychological disorders creates a shared language among mental-health providers and aids in clinical research. While diagnoses were recognized as far back as the Greeks, it was not until 1883 that German psychiatrist Emil Kräpelin (1856–1926) published a comprehensive system of psychological disorders that centered around a pattern of symptoms (i.e., **syndrome**) suggestive of an underlying physiological cause. Other clinicians also suggested popular classification systems but the need for a single, shared system paved the way for the American Psychiatric Association's 1952 publication of the first *Diagnostic and Statistical Manual* (DSM).

The DSM has undergone various revisions (in 1968, 1980, 1987, 1994, 2000, 2013), and it is the 1980 DSM-III version that began a multiaxial classification system that took into account the entire individual rather than just the specific problem behavior. Axes I and II contain the clinical diagnoses, including intellectual disability and personality disorders. Axes III and IV list any relevant medical conditions or psychosocial or environmental stressors, respectively. Axis V provides a global assessment of the individual's level of functioning. The most recent version -- the DSM-5-- has combined the first three axes and removed the last two. These revisions reflect an attempt to help clinicians streamline diagnosis and work better with other diagnostic systems such as health diagnoses outlined by the World Health Organization.

While the DSM has provided a necessary shared language for clinicians, aided in clinical research, and allowed clinicians to be reimbursed by insurance companies for their services, it is not without criticism. The DSM is based on clinical and research findings from Western culture, primarily the United States. It is also a medicalized categorical classification system that assumes disordered behavior does not differ in degree but in kind, as opposed to a dimensional classification system that would plot disordered behavior along a continuum. Finally, the number of diagnosable disorders has tripled since it was first published in 1952, so that almost half of Americans will have a diagnosable disorder in their lifetime, contributing to the continued concern of labeling and stigmatizing mentally ill individuals. These concerns appear to be relevant even in the DSM-5 version that came out in May of 2013.

Outside Resources

Video: An introduction to and overview of psychology, from its origins in the nineteenth century to current study of the brain\'s biochemistry.
http://www.learner.org/series/discoveringpsychology/01/e01expand.html

Video: The BBC provides an overview of ancient Greek approaches to health and medicine.
https://www.tes.com/teaching-resource/ancient-greek-approaches-to-health-and-medicine-6176019

Web: Images from the History of Medicine. Search \\\"mental illness\\\"
http://ihm.nlm.nih.gov/luna/servlet/view/all

Web: Science Museum Brought to Life
http://www.sciencemuseum.org.uk/broughttolife/themes/menalhealthandillness.aspx

Web: The Social Psychology Network provides a number of links and resources.
https://www.socialpsychology.org/history.htm

Web: The Wellcome Library. Search \\\"mental illness\\\".
http://wellcomelibrary.org/

Web: UCL Department of Science and Technology Studies
https://www.ucl.ac.uk/sts/

Web: US National Library of Medicine
http://vsearch.nlm.nih.gov/vivisimo/cgi-bin/query-meta?query=mental+illness&v:project=nlm-main-website

Discussion Questions

1. What does it mean to say that someone is mentally ill? What criteria are usually considered to determine whether someone is mentally ill?
2. Describe the difference between supernatural, somatogenic, and psychogenic theories of mental illness and how subscribing to a particular etiological theory determines the type of treatment used.
3. How did the Greeks describe hysteria and what treatment did they prescribe?
4. Describe humorism and how it explained mental illness.

5. Describe how the witch hunts came about and their relationship to mental illness.
6. Describe the development of treatment facilities for the mentally insane, from asylums to community mental health centers.
7. Describe the humane treatment of the mentally ill brought about by Chiarughi, Pinel, and Tuke in the late 18th and early 19th centuries and how it differed from the care provided in the centuries preceding it.
8. Describe William Tuke's treatment of the mentally ill at the York Retreat within the context of the Quaker Society of Friends. What influence did Tuke's treatment have in other parts of the world?
9. What are the 20th-century treatments resulting from the psychogenic and somatogenic theories of mental illness?
10. Describe why a classification system is important and how the leading classification system used in the United States works. Describe some concerns with regard to this system.

Vocabulary

Animism
The belief that everyone and everything had a “soul” and that mental illness was due to animistic causes, for example, evil spirits controlling an individual and his/her behavior.

Asylum
A place of refuge or safety established to confine and care for the mentally ill; forerunners of the mental hospital or psychiatric facility.

Biopsychosocial model
A model in which the interaction of biological, psychological, and sociocultural factors is seen as influencing the development of the individual.

Cathartic method
A therapeutic procedure introduced by Breuer and developed further by Freud in the late 19th century whereby a patient gains insight and emotional relief from recalling and reliving traumatic events.

Cultural relativism
The idea that cultural norms and values of a society can only be understood on their own terms or in their own context.

Etiology
The causal description of all of the factors that contribute to the development of a disorder or illness.

Humorism (or humoralism)
A belief held by ancient Greek and Roman physicians (and until the 19th century) that an excess or deficiency in any of the four bodily fluids, or humors—blood, black bile, yellow bile, and phlegm—directly affected their health and temperament.

Hysteria
Term used by the ancient Greeks and Egyptians to describe a disorder believed to be caused by a woman’s uterus wandering throughout the body and interfering with other organs (today referred to as conversion disorder, in which psychological problems are expressed in physical form).

Maladaptive
Term referring to behaviors that cause people who have them physical or emotional harm, prevent them from functioning in daily life, and/or indicate that they have lost touch with reality and/or cannot control their thoughts and behavior (also called dysfunctional).

Mesmerism
Derived from Franz Anton Mesmer in the late 18th century, an early version of hypnotism in which Mesmer claimed that hysterical symptoms could be treated through animal magnetism emanating from Mesmer's body and permeating the universe (and later through magnets); later explained in terms of high suggestibility in individuals.

Psychogenesis
Developing from psychological origins.

Somatogenesis
Developing from physical/bodily origins.

Supernatural
Developing from origins beyond the visible observable universe.

Syndrome
Involving a particular group of signs and symptoms.

"Traitement moral" (moral treatment)
A therapeutic regimen of improved nutrition, living conditions, and rewards for productive behavior that has been attributed to Philippe Pinel during the French Revolution, when he released mentally ill patients from their restraints and treated them with compassion and dignity rather than with contempt and denigration.

Trephination
The drilling of a hole in the skull, presumably as a way of treating psychological disorders.

References

Bell, L. V. (1980). *Treating the mentally ill: From colonial times to the present*. New York: Praeger.

Forrest, D. (1999). *Hypnotism: A history*. New York: Penguin.

Grob, G. N. (1994). *The mad among us: A history of the care of America's mentally ill*. New York: Free Press.

Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., . . . Krause, E. D. (2002). The dodo bird verdict is alive and well—mostly. *Clinical Psychology: Science and Practice, 9*, 2–12.

Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science and Practice, 9*(1), 21–25.

Micale, M. S. (1985). The Salpêtrière in the age of Charcot: An institutional perspective on medical history in the late nineteenth century. *Journal of Contemporary History, 20*, 703–731.

Restak, R. (2000). *Mysteries of the mind*. Washington, DC: National Geographic Society.

Schoeneman, T. J. (1977). The role of mental illness in the European witch hunts of the sixteenth and seventeenth centuries: An assessment. *Journal of the History of the Behavioral Sciences, 13*(4), 337–351.

Tseng, W. (1973). The development of psychiatric concepts in traditional Chinese medicine. *Archives of General Psychiatry, 29*, 569–575.

Viney, W., & Zorich, S. (1982). Contributions to the history of psychology: XXIX. Dorothea Dix and the history of psychology. *Psychological Reports, 50*, 211–218.

Zilboorg, G., & Henry, G. W. (1941). *A history of medical psychology*. New York: W. W. Norton.

37
Mood Disorders

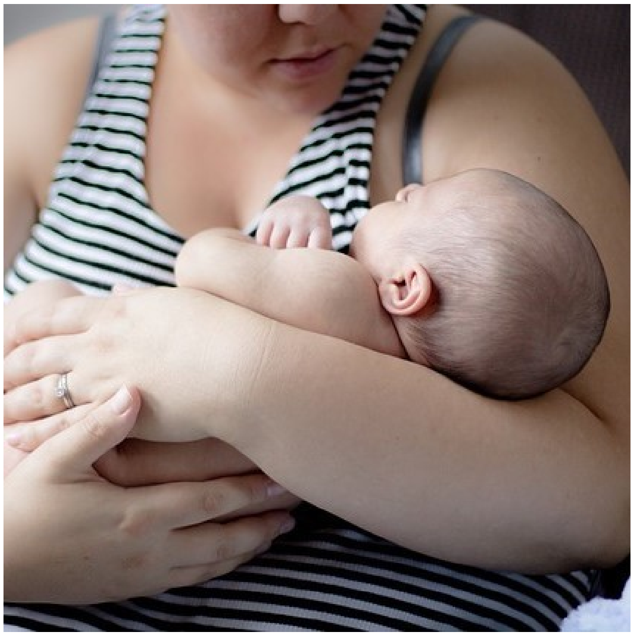
Anda Gershon & Renee Thompson

Everyone feels down or euphoric from time to time, but this is different from having a mood disorder such as major depressive disorder or bipolar disorder. Mood disorders are extended periods of depressed, euphoric, or irritable moods that in combination with other symptoms cause the person significant distress and interfere with his or her daily life, often resulting in social and occupational difficulties. In this module, we describe major mood disorders, including their symptom presentations, general prevalence rates, and how and why the rates of these disorders tend to vary by age, gender, and race. In addition, biological and environmental risk factors that have been implicated in the development and course of mood disorders, such as heritability and stressful life events, are reviewed. Finally, we provide an overview of treatments for mood disorders, covering treatments with demonstrated effectiveness, as well as new treatment options showing promise.

Learning Objectives

- Describe the diagnostic criteria for mood disorders.
- Understand age, gender, and ethnic differences in prevalence rates of mood disorders.
- Identify common risk factors for mood disorders.
- Know effective treatments of mood disorders.

The actress Brooke Shields published a memoir titled Down Came the Rain: My Journey through Postpartum Depression in which she described her struggles with depression following the birth of her daughter. Despite the fact that about one in 20 women experience



Perinatal depression following child birth afflicts about 5% of all mothers. An unfortunate social stigma regarding this form of depression compounds the problem for the women who suffer its effects. [Image: CC0 Public Domain]

depression after the birth of a baby (American Psychiatric Association [APA], 2013), postpartum depression—recently renamed "perinatal depression"—continues to be veiled by stigma, owing in part to a widely held expectation that motherhood should be a time of great joy. In an opinion piece in the *New York Times*, Shields revealed that entering motherhood was a profoundly overwhelming experience for her. She vividly describes experiencing a sense of "doom" and "dread" in response to her newborn baby. Because motherhood is conventionally thought of as a joyous event and not associated with sadness and hopelessness, responding to a newborn baby in this way can be shocking to the new mother as well as those close to her. It may also involve a great deal of shame for the mother, making her reluctant to divulge her experience to others, including her doctors and family.

Feelings of shame are not unique to perinatal depression. Stigma applies to other types of depressive and bipolar disorders and contributes to people not always receiving the necessary support and treatment for these disorders. In fact, the World Health Organization ranks both major depressive disorder (MDD) and bipolar disorder (BD) among the top 10 leading causes of disability worldwide. Further, MDD and BD carry a high risk of suicide. It is estimated that 25%–50% of people diagnosed with BD will attempt suicide at least once in their lifetimes (Goodwin & Jamison, 2007).

What Are Mood Disorders?

Mood Episodes

Everyone experiences brief periods of sadness, irritability, or euphoria. This is different than having a mood disorder, such as MDD or BD, which are characterized by a constellation of symptoms that causes people significant distress or impairs their everyday functioning.

Major Depressive Episode

A major depressive episode (MDE) refers to symptoms that co-occur for at least two weeks and cause significant distress or impairment in functioning, such as interfering with work, school, or relationships. Core symptoms include feeling down or depressed or experiencing anhedonia—loss of interest or pleasure in things that one typically enjoys. According to the fifth edition of the *Diagnostic and Statistical Manual* (*DSM-5;* APA, 2013), the criteria for an MDE require five or more of the following nine symptoms, including one or both of the first two symptoms, for most of the day, nearly every day:

1. depressed mood
2. diminished interest or pleasure in almost all activities
3. significant weight loss or gain or an increase or decrease in appetite
4. insomnia or hypersomnia
5. psychomotor agitation or retardation
6. fatigue or loss of energy
7. feeling worthless or excessive or inappropriate guilt
8. diminished ability to concentrate or indecisiveness
9. recurrent thoughts of death, suicidal ideation, or a suicide attempt

These symptoms cannot be caused by physiological effects of a substance or a general medical condition (e.g., hypothyroidism).

Manic or Hypomanic Episode

The core criterion for a manic or hypomanic episode is a distinct period of abnormally and persistently euphoric, expansive, or irritable mood and persistently increased goal-directed activity or energy. The mood disturbance must be present for one week or longer in mania (unless hospitalization is required) or four days or longer in hypomania. Concurrently, at least three of the following symptoms must be present in the context of euphoric mood (or at least four in the context of irritable mood):

1. inflated self-esteem or grandiosity
2. increased goal-directed activity or psychomotor agitation

3. reduced need for sleep
4. racing thoughts or flight of ideas
5. distractibility
6. increased talkativeness
7. excessive involvement in risky behaviors

Manic episodes are distinguished from hypomanic episodes by their duration and associated impairment; whereas manic episodes must last one week and are defined by a significant impairment in functioning, hypomanic episodes are shorter and not necessarily accompanied by impairment in functioning.

Mood Disorders

Unipolar Mood Disorders

Two major types of unipolar disorders described by the *DSM-5* (APA, 2013) are major depressive disorder and persistent depressive disorder (PDD; dysthymia). MDD is defined by one or more MDEs, but no history of manic or hypomanic episodes. Criteria for PDD are feeling depressed most of the day for more days than not, for at least two years. At least two of the following symptoms are also required to meet criteria for PDD:

1. poor appetite or overeating
2. insomnia or hypersomnia
3. low energy or fatigue
4. low self-esteem
5. poor concentration or difficulty making decisions
6. feelings of hopelessness

Like MDD, these symptoms need to cause significant distress or impairment and cannot be due to the effects of a substance or a general medical condition. To meet criteria for PDD, a person cannot be without symptoms for more than two months at a time. PDD has overlapping symptoms with MDD. If someone meets criteria for an MDE during a PDD episode, the person will receive diagnoses of PDD and MDD.

Bipolar Mood Disorders

Bipolar disorders are characterized by cycles of high energy and depression. [Image: Brett Whaley, https://goo.gl/k4HTR7, CC BY-NC 2.0, https://goo.gl/VnKlK8]

Three major types of BDs are described by the *DSM-5* (APA, 2013). Bipolar I Disorder (BD I), which was previously known as manic-depression, is characterized by a single (or recurrent) manic episode. A depressive episode is not necessary but commonly present for the diagnosis of BD I. Bipolar II Disorder is characterized by single (or recurrent) hypomanic episodes and depressive episodes. Another type of BD is cyclothymic disorder, characterized by numerous and alternating periods of hypomania and depression, lasting at least two years. To qualify for cyclothymic disorder, the periods of depression cannot meet full diagnostic criteria for an MDE; the person must experience symptoms at least half the time with no more than two consecutive symptom-free months; and the symptoms must cause significant distress or impairment.

It is important to note that the *DSM-5* was published in 2013, and findings based on the updated manual will be forthcoming. Consequently, the research presented below was largely based on a similar, but not identical, conceptualization of mood disorders drawn from the *DSM-IV* (APA, 2000).

How Common Are Mood Disorders? Who Develops Mood Disorders?

Depressive Disorders

In a nationally representative sample, lifetime prevalence rate for MDD is 16.6% (Kessler, Berglund, Demler, Jin, Merikangas, & Walters, 2005). This means that nearly one in five Americans will meet the criteria for MDD during their lifetime. The 12-month prevalence—the proportion of people who meet criteria for a disorder during a 12-month period—for PDD is approximately 0.5% (APA, 2013).

Although the onset of MDD can occur at any time throughout the lifespan, the average age of onset is mid-20s, with the age of onset decreasing with people born more recently (APA, 2000). Prevalence of MDD among older adults is much lower than it is for younger cohorts (Kessler, Birnbaum, Bromet, Hwang, Sampson, & Shahly, 2010). The duration of MDEs varies widely. Recovery begins within three months for 40% of people with MDD and within 12 months for 80% (APA, 2013). MDD tends to be a recurrent disorder with about 40%–50% of those who experience one MDE experiencing a second MDE (Monroe & Harkness, 2011). An earlier age of onset predicts a worse course. About 5%–10% of people who experience an MDE will later experience a manic episode (APA, 2000), thus no longer meeting criteria for MDD but instead meeting them for BD I. Diagnoses of other disorders across the lifetime are common for people with MDD: 59% experience an anxiety disorder; 32% experience an impulse control disorder, and 24% experience a substance use disorder (Kessler, Merikangas, & Wang, 2007).

Box 1. Specifiers

Both MDEs and manic episodes can be further described using standardized tags based on the timing of, or other symptoms that are occurring during, the mood episode, to increase diagnostic specificity and inform treatment. Psychotic features is specified when the episodes are accompanied by delusions (rigidly held beliefs that are false) or hallucinations (perceptual disturbances that are not based in reality). Seasonal pattern is specified when a mood episode occurs at the same time of the year for two consecutive years—most commonly occurring in the fall and winter. Peripartum onset is specified when a mood episode has an onset during pregnancy or within four weeks of the birth of a child. Approximately 3%–6% of women who have a child experience an MDE with peripartum onset (APA, 2013). This is less frequent and different from the baby blues or when women feel transient mood symptoms usually within 10 days of giving birth, which are experienced by most women (Nolen-Hoeksema & Hilt, 2009).

Women experience two to three times higher rates of MDD than do men (Nolen-Hoeksema & Hilt, 2009). This gender difference emerges during puberty (Conley & Rudolph, 2009). Before puberty, boys exhibit similar or higher prevalence rates of MDD than do girls (Twenge & Nolen-Hoeksema, 2002). MDD is inversely correlated with socioeconomic status (SES), a person's economic and social position based on income, education, and occupation. Higher prevalence rates of MDD are associated with lower SES (Lorant, Deliege, Eaton, Robert, Philippot, & Ansseau, 2003), particularly for adults over 65 years old (Kessler et al., 2010). Independent of SES, results from a nationally representative sample found that European Americans had a higher prevalence rate of MDD than did African Americans and Hispanic Americans, whose rates were similar (Breslau, Aguilar-Gaxiola, Kendler, Su, Williams, & Kessler, 2006). The course of MDD for African Americans is often more severe and less often treated than it is for European Americans, however (Williams et al., 2007). Native Americans have a higher prevalence rate than do European Americans, African Americans, or Hispanic Americans (Hasin, Goodwin, Stinson & Grant, 2005). Depression is not limited to industrialized or western cultures; it is found in all countries that have been examined, although the symptom presentation as well as prevalence rates vary across cultures (Chentsova-Dutton & Tsai, 2009).

Bipolar Disorders

Adolescents experience a higher incidence of bipolar spectrum disorders than do adults. Making matters worse, those who are diagnosed with BD at a younger age seem to suffer symptoms more intensely than those with adult onset. [Image: CC0 Public Domain]

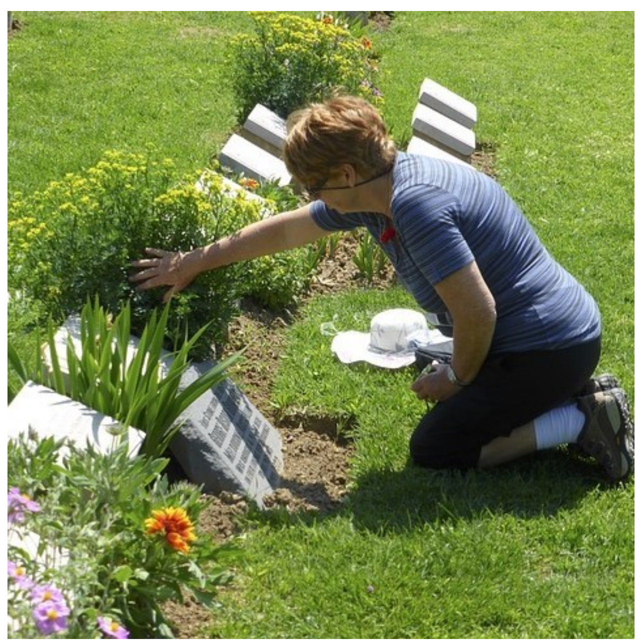
The lifetime prevalence rate of bipolar spectrum disorders in the general U.S. population is estimated at approximately 4.4%, with BD I constituting about 1% of this rate (Merikangas et al., 2007). Prevalence estimates, however, are highly dependent on the diagnostic procedures used (e.g., interviews vs. self-report) and whether or not sub-threshold forms of the disorder are included in the estimate. BD often co-occurs with other psychiatric disorders. Approximately 65% of people with BD meet diagnostic criteria for at least one additional psychiatric disorder, most commonly anxiety disorders and substance use disorders (McElroy et al., 2001). The co-occurrence of BD with other psychiatric disorders is associated with poorer illness course, including higher rates of suicidality (Leverich et al., 2003). A recent cross-national study sample of more than 60,000 adults from 11 countries, estimated the worldwide prevalence of BD at 2.4%, with BD I constituting 0.6% of this rate (Merikangas et al., 2011). In this study, the prevalence of BD varied somewhat by country. Whereas the United States had the highest lifetime prevalence (4.4%), India had the lowest (0.1%). Variation in prevalence rates was not necessarily related to SES, as in the case of Japan, a high-income country with a very low prevalence rate of BD (0.7%).

With regard to ethnicity, data from studies not confounded by SES or inaccuracies in diagnosis are limited, but available reports suggest rates of BD among European Americans are similar to those found among African Americans (Blazer et al., 1985) and Hispanic Americans (Breslau, Kendler, Su, Gaxiola-Aguilar, & Kessler, 2005). Another large community-based study found that although prevalence rates of mood disorders were similar across ethnic groups, Hispanic Americans and African Americans with a mood disorder were more likely to remain persistently ill than European Americans (Breslau et al., 2005). Compared with European Americans with BD, African Americans tend to be underdiagnosed for BD (and over-diagnosed for schizophrenia) (Kilbourne, Haas, Mulsant, Bauer, & Pincus, 2004; Minsky, Vega, Miskimen, Gara, & Escobar, 2003), and Hispanic Americans with BD have been shown to receive fewer psychiatric medication prescriptions and specialty treatment visits (Gonzalez et al., 2007). Misdiagnosis of BD can result in the underutilization of treatment or the utilization of inappropriate treatment, and thus profoundly impact the course of illness.

As with MDD, adolescence is known to be a significant risk period for BD; mood symptoms start by adolescence in roughly half of BD cases (Leverich et al., 2007; Perlis et al., 2004*).* Longitudinal studies show that those diagnosed with BD prior to adulthood experience a more pernicious course of illness relative to those with adult onset, including more episode recurrence, higher rates of suicidality, and profound social, occupational, and economic repercussions (e.g., Lewinsohn, Seeley, Buckley, & Klein, 2002). The prevalence of BD is substantially lower in older adults compared with younger adults (1% vs. 4%) (Merikangas et al., 2007).

What Are Some of the Factors Implicated in the Development and Course of Mood Disorders?

Mood disorders are complex disorders resulting from multiple factors. Causal explanations can be attempted at various levels, including biological and psychosocial levels. Below are several of the key factors that contribute to onset and course of mood disorders are highlighted.



Romantic relationships can affect mood as in the case of divorce or the death of a spouse. [Image: CC0 Public Domain]

Depressive Disorders

Research across family and twin studies has provided support that genetic factors are implicated in the development of MDD. Twin studies suggest that familial influence on MDD is mostly due to genetic effects and that individual-specific environmental effects (e.g., romantic relationships) play an important role, too. By contrast, the contribution of shared environmental effect by siblings is negligible (Sullivan, Neale & Kendler, 2000). The mode of inheritance is not fully understood although no single genetic variation has been found to increase the risk of MDD significantly. Instead, several genetic variants and environmental factors most likely contribute to the risk for MDD (Lohoff, 2010).

One environmental stressor that has received much support in relation to MDD is stressful life events. In particular, severe stressful life events—those that have long-term consequences and involve loss of a significant relationship (e.g., divorce) or economic stability (e.g., unemployment) are strongly related to depression (Brown & Harris, 1989; Monroe et al., 2009). Stressful life events are more likely to predict the first MDE than subsequent episodes (Lewinsohn, Allen, Seeley, & Gotlib, 1999). In contrast, minor events may play a larger role in subsequent episodes than the initial episodes (Monroe & Harkness, 2005).

Depression research has not been limited to examining reactivity to stressful life events. Much research, particularly brain imagining research using functional magnetic resonance imaging (fMRI), has centered on examining neural circuitry—the interconnections that allow multiple brain regions to perceive, generate, and encode information in concert. A meta-analysis of neuroimaging studies showed that when viewing negative stimuli (e.g., picture of an angry face, picture of a car accident), compared with healthy control participants, participants with MDD have greater activation in brain regions involved in stress response and reduced activation of brain regions involved in positively motivated behaviors (Hamilton, Etkin, Furman, Lemus, Johnson, & Gotlib, 2012).

Other environmental factors related to increased risk for MDD include experiencing **early adversity** (e.g., childhood abuse or neglect; Widom, DuMont, & Czaja, 2007), **chronic stress** (e.g., poverty) and interpersonal factors. For example, marital dissatisfaction predicts increases in depressive symptoms in both men and women. On the other hand, depressive symptoms also predict increases in marital dissatisfaction (Whisman & Uebelacker, 2009). Research has found that people with MDD generate some of their interpersonal stress (Hammen, 2005). People with MDD whose relatives or spouses can be described as critical and emotionally overinvolved have higher relapse rates than do those living with people who are less critical and emotionally overinvolved (Butzlaff & Hooley, 1998).

People's **attributional styles** or their general ways of thinking, interpreting, and recalling information have also been examined in the etiology of MDD (Gotlib & Joormann, 2010). People with a pessimistic attributional style tend to make internal (versus external), global (versus specific), and stable (versus unstable) attributions to negative events, serving as a vulnerability to developing MDD. For example, someone who when he fails an exam thinks that it was his fault (internal), that he is stupid (global), and that he will always do poorly (stable) has a pessimistic attribution style. Several influential theories of depression incorporate attributional styles (Abramson, Metalsky, & Alloy, 1989; Abramson Seligman, & Teasdale, 1978).

Bipolar Disorders

Although there have been important advances in research on the etiology, course, and treatment of BD, there remains a need to understand the mechanisms that contribute to episode onset and relapse. There is compelling evidence for biological causes of BD, which is known to be highly heritable (McGuffin, Rijsdijk, Andrew, Sham, Katz, & Cardno, 2003). It may be argued that a high rate of heritability demonstrates that BD is fundamentally a biological phenomenon. However, there is much variability in the course of BD both within a person across time and across people (Johnson, 2005). The triggers that determine how and when this genetic vulnerability is expressed are not yet understood; however, there is evidence to suggest that psychosocial triggers may play an important role in BD risk (e.g., Johnson et al., 2008; Malkoff-Schwartz et al., 1998).

In addition to the genetic contribution, biological explanations of BD have also focused on brain function. Many of the studies using fMRI techniques to characterize BD have focused on the processing of emotional stimuli based on the idea that BD is fundamentally a disorder of emotion (APA, 2000). Findings show that regions of the brain thought to be involved in emotional processing and regulation are activated differently in people with BD relative to healthy controls (e.g., Altshuler et al., 2008; Hassel et al., 2008; Lennox, Jacob, Calder, Lupson, & Bullmore, 2004).

However, there is little consensus as to whether a particular brain region becomes more or less active in response to an emotional stimulus among people with BD compared with healthy controls. Mixed findings are in part due to samples consisting of participants who are at various phases of illness at the time of testing (manic, depressed, inter-episode). Sample sizes tend to be relatively small, making comparisons between subgroups difficult. Additionally, the use of a standardized stimulus (e.g., facial expression of anger) may not elicit a sufficiently strong response. Personally engaging stimuli, such as recalling a memory, may be more effective in inducing strong emotions (Isacowitz, Gershon, Allard, & Johnson, 2013).

Within the psychosocial level, research has focused on the environmental contributors to BD. A series of studies show that environmental stressors, particularly severe stressors (e.g., loss of a significant relationship), can adversely impact the course of BD. People with BD have substantially increased risk of relapse (Ellicott, Hammen, Gitlin, Brown, & Jamison, 1990) and suffer more depressive symptoms (Johnson, Winett, Meyer, Greenhouse, & Miller, 1999) following a severe life stressor. Interestingly, positive life events can also adversely impact the course of BD. People with BD suffer more manic symptoms after life events involving attainment of a desired goal (Johnson et al., 2008). Such findings suggest that people with BD may have a hypersensitivity to rewards.

Evidence from the life stress literature has also suggested that people with mood disorders may have a circadian vulnerability that renders them sensitive to stressors that disrupt their sleep or rhythms. According to social zeitgeber theory (Ehlers, Frank, & Kupfer, 1988; Frank et al., 1994), stressors that disrupt sleep, or that disrupt the daily routines that entrain the biological clock (e.g., meal times) can trigger episode relapse. Consistent with this theory, studies have shown that life events that involve a disruption in sleep and daily routines, such as overnight travel, can increase bipolar symptoms in people with BD (Malkoff-Schwartz et al., 1998).

What Are Some of the Well-Supported Treatments for Mood Disorders?

Depressive Disorders

A number of medications are effective in treating mood disorders. Meditation, exercise, counseling and other therapies also show effectiveness for some disorders. [Image: CC0 Public Domain]

There are many treatment options available for people with MDD. First, a number of antidepressant medications are available, all of which target one or more of the neurotransmitters implicated in depression.The earliest antidepressant medications were monoamine oxidase inhibitors (MAOIs). MAOIs inhibit monoamine oxidase, an enzyme involved in deactivating dopamine, norepinephrine, and serotonin. Although effective in treating depression, MAOIs can have serious side effects. Patients taking MAOIs may develop dangerously high blood pressure if they take certain drugs (e.g., antihistamines) or eat foods containing tyramine, an amino acid commonly found in foods such as aged cheeses, wine, and soy sauce. Tricyclics, the second-oldest class of antidepressant medications, block the reabsorption of norepinephrine, serotonin, or dopamine at synapses, resulting in their increased availability. Tricyclics are most effective for treating vegetative and somatic symptoms of depression. Like MAOIs, they have serious side effects, the most concerning of which is being cardiotoxic. Selective serotonin reuptake inhibitors (SSRIs; e.g., Fluoxetine) and serotonin and norepinephrine reuptake inhibitors (SNRIs; e.g., Duloxetine) are the most recently introduced antidepressant medications. SSRIs, the most commonly prescribed antidepressant medication, block the reabsorption of serotonin, whereas SNRIs block the reabsorption of serotonin and norepinephrine. SSRIs and SNRIs have fewer serious side effects than do MAOIs and tricyclics. In particular, they are less cardiotoxic, less lethal in overdose, and produce fewer cognitive impairments. They are not, however, without their own side effects, which include but are not limited to difficulty having orgasms, gastrointestinal issues, and insomnia. It should be noted that anti-depressant medication may not work equally for all people. This approach to treatment often involves experimentation with several medications and dosages, and may be more effective when paired with physical exercise and psychotherapy.

Other biological treatments for people with depression include electroconvulsive therapy (ECT), transcranial magnetic stimulation (TMS), and deep brain stimulation. ECT involves inducing a seizure after a patient takes muscle relaxants and is under general anesthesia. ECT is viable treatment for patients with severe depression or who show resistance to antidepressants although the mechanisms through which it works remain unknown. A common side effect is confusion and memory loss, usually short-term (Schulze-Rauschenbach, Harms, Schlaepfer, Maier, Falkai, & Wagner, 2005). Repetitive TMS is a noninvasive technique administered while a patient is awake. Brief pulsating magnetic fields are delivered to the cortex, inducing electrical activity. TMS has fewer side effects than ECT (Schulze-Rauschenbach et al., 2005), and while outcome studies are mixed, there is evidence that TMS is a promising treatment for patients with MDD who have shown resistance to other treatments (Rosa et al., 2006). Most recently, deep brain stimulation is being examined as a treatment option for patients who did not respond to more traditional treatments like those already described. Deep brain stimulation involves implanting an electrode in the brain. The electrode is connected to an implanted neurostimulator, which electrically stimulates that particular brain region. Although there is some evidence of its effectiveness (Mayberg et al., 2005), additional research is needed.

Several psychosocial treatments have received strong empirical support, meaning that independent investigations have achieved similarly positive results—a high threshold for examining treatment outcomes. These treatments include but are not limited to behavior therapy, cognitive therapy, and interpersonal therapy. Behavior therapies focus on increasing the frequency and quality of experiences that are pleasant or help the patient achieve mastery.

Cognitive therapies primarily focus on helping patients identify and change distorted automatic thoughts and assumptions (e.g., Beck, 1967). Cognitive-behavioral therapies are based on the rationale that thoughts, behaviors, and emotions affect and are affected by each other. Interpersonal Therapy for Depression focuses largely on improving interpersonal relationships by targeting problem areas, specifically unresolved grief, interpersonal role disputes, role transitions, and interpersonal deficits. Finally, there is also some support for the effectiveness of Short-Term Psychodynamic Therapy for Depression (Leichsenring, 2001). The short-term treatment focuses on a limited number of important issues, and the therapist tends to be more actively involved than in more traditional psychodynamic therapy.

Bipolar Disorders

Patients with BD are typically treated with pharmacotherapy. Antidepressants such as SSRIs and SNRIs are the primary choice of treatment for depression, whereas for BD, lithium is the first line treatment choice. This is because SSRIs and SNRIs have the potential to induce mania or hypomania in patients with BD. Lithium acts on several neurotransmitter systems in the brain through complex mechanisms, including reduction of excitatory (dopamine and glutamate) neurotransmission, and increasing of inhibitory (GABA) neurotransmission (Lenox & Hahn, 2000). Lithium has strong efficacy for the treatment of BD (Geddes, Burgess, Hawton, Jamison, & Goodwin, 2004). However, a number of side effects can make lithium treatment difficult for patients to tolerate. Side effects include impaired cognitive function (Wingo, Wingo, Harvey, & Baldessarini, 2009), as well as physical symptoms such as nausea, tremor, weight gain, and fatigue (Dunner, 2000). Some of these side effects can improve with continued use; however, medication noncompliance remains an ongoing concern in the treatment of patients with BD. Anticonvulsant medications (e.g., carbamazepine, valproate) are also commonly used to treat patients with BD, either alone or in conjunction with lithium.

There are several adjunctive treatment options for people with BD. Interpersonal and social rhythm therapy (IPSRT; Frank et al., 1994) is a psychosocial intervention focused on addressing the mechanism of action posited in social *zeitgeber* theory to predispose patients who have BD to relapse, namely sleep disruption. A growing body of literature provides support for the central role of sleep dysregulation in BD (Harvey, 2008). Consistent with this literature, IPSRT aims to increase rhythmicity of patients' lives and encourage vigilance in maintaining a stable rhythm. The therapist and patient work to develop and maintain a healthy balance of activity and stimulation such that the patient does not become overly active (e.g., by taking on too many projects) or inactive (e.g., by avoiding social contact). The efficacy of IPSRT has been demonstrated in that patients who received this treatment show reduced risk of episode recurrence and are more likely to remain well (Frank et al., 2005).

Conclusion

Everyone feels down or euphoric from time to time. For some people, these feelings can last for long periods of time and can also co-occur with other symptoms that, in combination, interfere with their everyday lives. When people experience an MDE or a manic episode, they see the world differently. During an MDE, people often feel hopeless about the future, and may even experience suicidal thoughts. During a manic episode, people often behave in ways that are risky or place them in danger. They may spend money excessively or have unprotected sex, often expressing deep shame over these decisions after the episode. MDD and BD cause significant problems for people at school, at work, and in their relationships and affect people regardless of gender, age, nationality, race, religion, or sexual orientation. If you or someone you know is suffering from a mood disorder, it is important to seek help. Effective treatments are available and continually improving. If you have an interest in mood disorders, there are many ways to contribute to their understanding, prevention, and treatment, whether by engaging in research or clinical work.

Outside Resources

Books: Recommended memoirs include A Memoir of Madness by William Styron (MDD); Noonday Demon: An Atlas of Depression by Andrew Solomon (MDD); and An Unquiet Mind: A Memoir of Moods and Madness by Kay Redfield (BD).

Web: Visit the Association for Behavioral and Cognitive Therapies to find a list of the recommended therapists and evidence-based treatments.
http://www.abct.org

Web: Visit the Depression and Bipolar Support Alliance for educational information and social support options.
http://www.dbsalliance.org/

Discussion Questions

1. What factors might explain the large gender difference in the prevalence rates of MDD?
2. Why might American ethnic minority groups experience more persistent BD than European Americans?
3. Why might the age of onset for MDD be decreasing over time?
4. Why might overnight travel constitute a potential risk for a person with BD?
5. What are some reasons positive life events may precede the occurrence of manic episode?

Vocabulary

Anhedonia
Loss of interest or pleasure in activities one previously found enjoyable or rewarding.

Attributional style
The tendency by which a person infers the cause or meaning of behaviors or events.

Chronic stress
Discrete or related problematic events and conditions which persist over time and result in prolonged activation of the biological and/or psychological stress response (e.g., unemployment, ongoing health difficulties, marital discord).

Early adversity
Single or multiple acute or chronic stressful events, which may be biological or psychological in nature (e.g., poverty, abuse, childhood illness or injury), occurring during childhood and resulting in a biological and/or psychological stress response.

Grandiosity
Inflated self-esteem or an exaggerated sense of self-importance and self-worth (e.g., believing one has special powers or superior abilities).

Hypersomnia
Excessive daytime sleepiness, including difficulty staying awake or napping, or prolonged sleep episodes.

Psychomotor agitation
Increased motor activity associated with restlessness, including physical actions (e.g., fidgeting, pacing, feet tapping, handwringing).

Psychomotor retardation
A slowing of physical activities in which routine activities (e.g., eating, brushing teeth) are performed in an unusually slow manner.

Social zeitgeber
Zeitgeber is German for “time giver.” Social zeitgebers are environmental cues, such as meal times and interactions with other people, that entrain biological rhythms and thus sleep-wake cycle regularity.

Socioeconomic status (SES)

A person's economic and social position based on income, education, and occupation.

Suicidal ideation

Recurring thoughts about suicide, including considering or planning for suicide, or preoccupation with suicide.

References

Abramson, L. Y, Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49–74. doi: 10.1037/0021-843X.87.1.49

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96*, 358–373. doi: 10.1037/0022-3514.56.3.431

Altshuler, L., Bookheimer, S., Townsend, J., Proenza, M. A., Sabb, F., Mintz, J., & Cohen, M. S. (2008). Regional brain changes in bipolar I depression: A functional magnetic resonance imaging study. *Bipolar Disorders, 10*, 708–717. doi: 10.1111/j.1399-5618.2008.00617.x

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York, NY: Hoeber.

Blazer, D., George, L. K., Landerman, R., Pennybacker, M., Melville, M. L., Woodbury, M., et al. (1985). Psychiatric disorders. A rural/urban comparison. *Archives of General Psychiatry, 42*, 651–656. PMID: 4015306. doi: 10.1001/archpsyc.1985.01790300013002

Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying race-ethnic differences in risk for psychiatric disorder in a US national sample. *Psychological Medicine, 36*, 57–68. doi: 10.1017/S0033291705006161

Breslau, J., Kendler, K. S., Su, M., Gaxiola-Aguilar, S., & Kessler, R. C. (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. *Psychological Medicine, 35*, 317–327. doi: 10.1017/S0033291704003514

Brown, G. W., & Harris, T. O. (1989). *Life events and illness*. New York, NY: Guilford Press.

Butzlaff, R. L., & Hooley, J. M. (1998). Expressed emotion and psychiatric relapse: A meta-analysis. *Archives of General Psychiatry, 55*, 547–552. doi: 10.1001/archpsyc.55.6.547

Chentsova-Dutton, Y. E., & Tsai, J. L. (2009). Understanding depression across cultures. In I. H. Gotlib & C.L. Hammen (Eds.), *Handbook of depression* (2nd ed., pp. 363–385). New York, NY: Guilford Press.

Conley, C. S., & Rudolph, K. D. (2009). The emerging sex difference in adolescent depression: Interacting contributions of puberty and peer stress. *Development and Psychopathology, 21*, 593–620. doi: 10.1017/S0954579409000327

Dunner, D. L. (2000). Optimizing lithium treatment. *Journal of Clinical Psychiatry, 61*(S9), 76–81.

Ehlers, C. L., Frank, E., & Kupfer, D. J. (1988). Social zeitgebers and biological rhythms: a unified approach to understanding the etiology of depression. *Archives of General Psychiatry, 45*, 948–952. doi: 10.1001/archpsyc.1988.01800340076012

Ellicott, A., Hammen, C., Gitlin, M., Brown, G., & Jamison, K. (1990). Life events and the course of bipolar disorder. *American Journal of Psychiatry, 147*, 1194–1198.

Frank, E., Kupfer, D. J., Ehlers, C. L., Monk, T., Cornes, C., Carter, S., et al. (1994). Interpersonal and social rhythm therapy for bipolar disorder: Integrating interpersonal and behavioral approaches. *Behavior Therapy, 17*, 143–149.

Frank, E., Kupfer, D. J., Thase, M. E., Mallinger, A. G., Swartz, H. A., Fagiolini, A. M., et al. (2005). Two-year outcomes for interpersonal and social rhythm therapy in individuals with bipolar I disorder. *Archives of General Psychiatry, 62*, 996–1004. doi: 10.1001/archpsyc.62.9.996

Geddes, J. R., Burgess, S., Hawton, K., Jamison, K., & Goodwin, G. M. (2004). Long-term lithium therapy for bipolar disorder: systematic review and meta-analysis of randomized controlled trials. *American Journal of Psychiatry, 161*, 217–222. doi: 10.1176/appi.ajp.161.2.217

Gonzalez, J. M., Perlick, D. A., Miklowitz, D. J., Kaczynski, R., Hernandez, M., Rosenheck, R. A., et al. (2007). Factors associated with stigma among caregivers of patients with bipolar disorder in the STEP-BD study. *Psychiatric Services, 58*, 41–48. doi: 10.1176/appi.ps.58.1.41

Goodwin, F. K., & Jamison, K. R. (2007). *Manic-depressive illness: Bipolar disorders and recurrent depression*. New York, NY: Oxford University Press.

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology, 6*, 285–312. doi: 10.1146/annurev.clinpsy.121208.131305

Hamilton, J. P., Etkin, A., Furman, D. F., Lemus, M. G., Johnson, R. F., & Gotlib, I. H. (2012). Functional neuroimaging of major depressive disorder: A meta-analysis and new integration of baseline activation and neural response data. *American Journal of Psychiatry, 169*, 693–703.

Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1*, 293–319. doi: 10.1146/annurev.clinpsy.1.102803.143938

Harvey, A. G. (2008). Sleep and Circadian Rhythms in Bipolar Disorder: Seeking synchrony, harmony and regulation. *American Journal of Psychiatry, 165*, 820–829. doi: 10.1176/appi.ajp.2008.08010098

Hasin, D. S., Goodwin, R. D., Sintson, F. S., & Grant, B. F. (2005). Epidemiology of major depressive disorder: Results from the National Epidemiological Survey on Alcoholism and Related Conditions. *Archives of General Psychiatry, 62*, 1097–1106. doi: 10.1001/archpsyc.62.10.1097

Hassel, S., Almeida, J. R., Kerr, N., Nau, S., Ladouceur, C. D., Fissell, K., et al. (2008). Elevated striatal and decreased dorsolateral prefrontal cortical activity in response to emotional stimuli in euthymic bipolar disorder: No associations with psychotropic medication load. *Bipolar Disorders, 10*, 916–927. doi: 10.1111/j.1399-5618.2008.00641.x

Isaacowitz, D. M., Gershon, A., Allard, E. S., & Johnson, S. L. (2013). Emotion in aging and bipolar disorder: Similarities, differences and lessons for further research. *Emotion Review, 5*, 312–320. doi: 10.1177/1754073912472244

Johnson, S. L. (2005). Mania and dysregulation in goal pursuit: A review. *Clinical Psychology Review, 25*, 241–262. doi: 10.1016/j.cpr.2004.11.002

Johnson, S. L., Cueller, A. K., Ruggero, C., Winett-Perlman, C., Goodnick, P., White, R., et al. (2008). Life events as predictors of mania and depression in bipolar I disorder. *Journal of Abnormal Psychology, 117*, 268–277. doi: 10.1037/0021-843X.117.2.268

Johnson, S. L., Winett, C. A., Meyer, B., Greenhouse, W. J., & Miller, I. (1999). Social support and the course of bipolar disorder. *Journal of Abnormal Psychology, 108*, 558–566. doi: 10.1037/0021-843X.108.4.558

Kessler, R. C., Berglund, P., Demler, O., Jim, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*, 593–602. doi: 10.1001/archpsyc.62.6.593

Kessler, R. C., Birnbaum, H., Bromet, E., Hwang, I., Sampson, N., & Shahly, V. (2010). Age differences in major depression: Results from the National Comorbidity Surveys Replication (NCS-R). *Psychological Medicine, 40*, 225–237. doi: 10.1017/S0033291709990213

Kessler, R. C., Merikangas, K. R., & Wang, P. S. (2007). Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the 21st century. *Annual Review of Clinical Psychology, 3*, 137–158. doi: 10.1146/annurev.clinpsy.3.022806.091444

Kilbourne, A. M., Haas, G. L., Mulsant, B. H., Bauer, M. S., & Pincus, H. A. (2004) Concurrent psychiatric diagnoses by age and race among persons with bipolar disorder. *Psychiatric Services, 55*, 931–933. doi: 10.1176/appi.ps.55.8.931

Leichsenring, F. (2001). Comparative effects of short-term psychodynamic psychotherapy and cognitive-behavioral therapy in depression: A meta-analytic approach. *Clinical Psychology Review, 21*, 401–419. doi: 10.1016/S0272-7358(99)00057-4

Lennox, B. R., Jacob, R., Calder, A. J., Lupson, V., & Bullmore, E. T. (2004). Behavioural and neurocognitive responses to sad facial affect are attenuated in patients with mania. *Psychological Medicine, 34*, 795–802. doi: 10.1017/S0033291704002557

Lenox, R. H., & Hahn C. G. (2000). Overview of the mechanism of action of lithium in the brain:

fifty-year update. *Journal of Clinical Psychiatry, 61* (S9), 5–15.

Leverich, G. S., Altshuler, L. L., Frye, M. A., Suppes, T., Keck, P. E. Jr, McElroy, S. L., et al. (2003). Factors associated with suicide attempts in 648 patients with bipolar disorder in the Stanley Foundation Bipolar Network. *Journal of Clinical Psychiatry, 64*, 506–515. doi: 10.4088/JCP.v64n0503

Leverich, G. S., Post, R. M., Keck, P. E. Jr, Altshuler, L. L., Frye, M. A., Kupka, R. W., et al. (2007). The poor prognosis of childhood-onset bipolar disorder. *Journal of Pediatrics, 150*, 485–490. PMID: 17452221. doi: 10.1016/j.jpeds.2006.10.070

Lewinsohn, P. M., Allen, N. B., Seeley, J. R., & Gotlib, I. H. (1999). First onset versus recurrence of depression: differential processes of psychosocial risk. *Journal of Abnormal Psychology, 108*, 483–489. doi: 10.1037/0021-843X.108.3.483

Lewinsohn, P. M., Seeley, J. R., Buckley, M. E., & Klein, D. N. (2002). Bipolar disorder in adolescence and young adulthood. *Child & Adolescent Psychiatric Clinics of North America, 11*, 461–475. doi: 10.1016/S1056-4993(02)00005-6

Lohoff, F. W. (2010). Overview of genetics of major depressive disorder. *Current Psychiatry Reports, 12*, 539–546. doi: 10.1007/s11920-010-0150-6

Lorant, V., Deliege, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, A. (2003). Socioeconomic inequalities in depression: A meta-analysis. *American Journal of Epidemiology, 157*, 98–112. doi: 10.1093/aje/kwf182

Malkoff-Schwartz, S., Frank, E., Anderson, B. P., Sherrill, J. T., Siegel, L., Patterson, D., et al. (1998). Stressful life events and social rhythm disruption in the onset of manic and depressive bipolar episodes: a preliminary investigation. *Archives of General Psychiatry, 55*, 702–707. doi: 10.1001/archpsyc.55.8.702

Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowixz, D., Hamani, C., Schwalb, J. M., & Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron, 45*, 651–660. doi: 10.1016/j.neuron.2005.02.014

McElroy, S. L., Altshuler, L. L., Suppes, T., Keck, P. E. Jr, Frye, M. A., Denicoff, K. D., et al. (2001). Axis I psychiatric comorbidity and its relationship to historical illness variables in 288 patients with bipolar disorder. *American Journal of Psychiatry, 158*, 420–426. doi: 10.1176/appi.ajp.158.3.420

McGuffin, P., Rijsdijk, F., Andrew, M., Sham, P., Katz, R., Cardno, A. (2003). The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Archives of General Psychiatry, 60*, 497–502. doi: 10.1001/archpsyc.60.5.497

Merikangas, K. R., Akiskal, H. S., Angst, J., Greenberg, P. E., Hirschfeld, R. M., Petukhova, M., et al. (2007). Lifetime and 12-month prevalence of bipolar spectrum disorder in the National

Comorbidity Survey replication. *Archives of General Psychiatry, 64*, 543–552. doi: 10.1001/archpsyc.64.5.543

Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., et al. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of General Psychiatry, 68*, 241–251. doi: 10.1001/archgenpsychiatry.2011.12

Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of General Psychiatry, 60*, 637–644. doi: 10.1001/archpsyc.60.6.637

Monroe, S. M., & Harkness, K. L. (2011). Recurrence in major depression: A conceptual analysis. *Psychological Review, 118*, 655–674. doi: 10.1037/a0025190

Monroe, S. M., & Harkness, K. L. (2005). Life stress, the "Kindling" hypothesis, and the recurrence of depression: Considerations from a life stress perspective. *Psychological Review, 112*, 417–445. doi: 10.1037/0033-295X.112.2.417

Monroe, S.M., Slavich, G.M., Georgiades, K. (2009). The social environment and life stress in depression. In Gotlib, I.H., Hammen, C.L (Eds.) *Handbook of depression* (2nd ed., pp. 340-360). New York, NY: Guilford Press.

Nolen-Hoeksema, S., & Hilt, L. M. (2009). Gender differences in depression. In I. H. Gotlib & Hammen, C. L. (Eds.), *Handbook of depression* (2nd ed., pp. 386–404). New York, NY: Guilford Press.

Perlis, R. H., Miyahara, S., Marangell, L. B., Wisniewski, S. R., Ostacher, M., DelBello, M. P., et al. (2004). Long-term implications of early onset in bipolar disorder: data from the first 1000 participants in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological Psychiatry, 55*, 875–881. PMID: 15110730. doi: 10.1016/j.pscychresns.2007.10.003

Rosa, M. A., Gattaz, W. F., Pascual-Leone, A., Fregni, F., Rosa, M. O., Rumi, D. O., … Marcolin, M. A. (2006). Comparison of repetitive transcranial magnetic stimulation and electroconvulsive therapy in unipolar non-psychotic refractory depression: a randomized, single-blind study. *International Journal of Neuropsychopharmacology, 9*, 667–676. doi: 10.1017/S1461145706007127

Schulze-Rauschenbach, S. C., Harms, U., Schlaepfer, T. E., Maier, W., Falkai, P., & Wagner, M. (2005). Distinctive neurocognitive effects of repetitive transcranial magnetic stimulation and electroconvulsive therapy in major depression. *British Journal of Psychiatry, 186*, 410–416. doi: 10.1192/bjp.186.5.410

Shields, B. (2005). *Down Came the Rain: My Journey Through Postpartum Depression*. New York: Hyperion.

Sullivan, P., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression:

Review and meta-analysis. *American Journal of Psychiatry, 157*, 1552–1562. doi: 10.1176/appi.ajp.157.10.1552

Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, SES, and birth cohort differences on the Children's Depression Inventory: A meta-analysis. *Journal of Abnormal Psychology, 111*, 578–588. doi: 10.1037/0021-843X.111.4.578

Whisman, M. A., & Uebelacker, L. A. (2009). Prospective associations between marital discord and depressive symptoms in middle-aged and older adults. *Psychology and Aging, 24*, 184–189. doi: 10.1037/a0014759

Widom, C. S., DuMont, K., & Czaja, S. J. (2007). A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. *Archives of General Psychiatry, 64*, 49–56. doi: 10.1001/archpsyc.64.1.49

Williams, D. R., Gonzalez, H. M., Neighbors, H., Nesse, R., Abelson, J. M., Sweetman, J., & Jackson, J. S. (2007). Prevalence and distribution of major depressive disorder in African Americans, Caribbean blacks, and non-Hispanic whites: Results from the National Survey of American Life. *Archives of General Psychiatry, 64*, 305–315. doi: 10.1001/archpsyc.64.3.305

Wingo, A. P., Wingo, T. S., Harvey, P. D., & Baldessarini, R. J. (2009). Effects of lithium on cognitive performance: a meta-analysis. *Journal of Clinical Psychiatry, 70*, 1588–1597. doi: 10.4088/JCP.08r04972

38
Autism: Insights from the Study of the Social Brain

Kevin A. Pelphrey

People with autism spectrum disorder (ASD) suffer from a profound social disability. Social neuroscience is the study of the parts of the brain that support social interactions or the "social brain." This module provides an overview of ASD and focuses on understanding how social brain dysfunction leads to ASD. Our increasing understanding of the social brain and its dysfunction in ASD will allow us to better identify the genes that cause ASD and will help us to create and pick out treatments to better match individuals. Because social brain systems emerge in infancy, social neuroscience can help us to figure out how to diagnose ASD even before the symptoms of ASD are clearly present. This is a hopeful time because social brain systems remain malleable well into adulthood and thus open to creative new interventions that are informed by state-of-the-art science.

Learning Objectives

- Know the basic symptoms of ASD.
- Distinguish components of the social brain and understand their dysfunction in ASD.
- Appreciate how social neuroscience may facilitate the diagnosis and treatment of ASD.

Defining Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a developmental disorder that usually emerges in the first three years and persists throughout the individual's life. Though the key symptoms of ASD fall into three general categories (see below), each person with ASD exhibits symptoms in these domains in different ways and to varying degrees. This *phenotypic heterogeneity* reflects the high degree of variability in the genes underlying ASD (Geschwind & Levitt, 2007). Though we have identified genetic differences associated with individual cases of ASD, each accounts for only a small number of the actual cases, suggesting that no single genetic cause will apply in the majority of people with ASD. There is currently no biological test for ASD.

Autism is in the category of *pervasive developmental disorders*, which includes Asperger's disorder, childhood disintegrative disorder, autistic disorder, and pervasive developmental disorder - not otherwise specified. These disorders, together, are labeled *autism spectrum disorder* (ASD). ASD is defined by the presence of profound difficulties in social interactions and communication combined with the presence of repetitive or restricted interests, cognitions and behaviors. The diagnostic process involves a combination of parental report and clinical observation. Children with significant impairments across the social/communication domain who also exhibit repetitive behaviors can qualify for the ASD diagnosis. There is wide variability in the precise symptom profile an individual may exhibit.

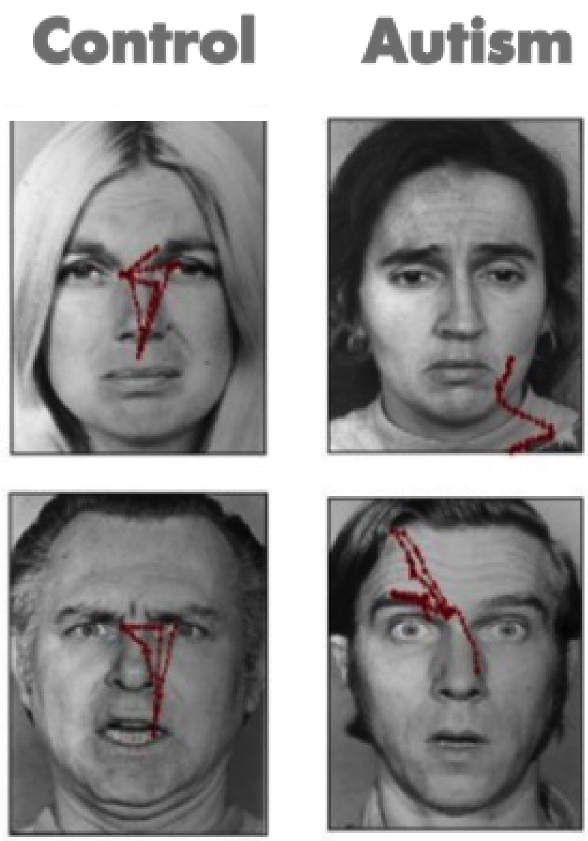


Figure 1. The red lines indicate the scan paths (collection of eye movements) used by people with (right column) and without (left column) autism to explore faces. Modified from Pelphrey et al., (2002).

Since Kanner first described ASD in 1943, important commonalities in symptom presentation have been used to compile criteria for the diagnosis of ASD. These diagnostic criteria have evolved during the past 70 years and continue to evolve (e.g., see the recent changes to the diagnostic criteria on the American Psychiatric Association's website, **http://www.dsm5.org/),** yet impaired social functioning remains a required symptom for an ASD diagnosis. Deficits in social functioning are present in varying degrees for simple behaviors such as eye contact, and complex behaviors like navigating the give and take of a group conversation for individuals of all functioning levels (i.e. high or low IQ). Moreover, difficulties with social information processing occur in both visual (e.g., Pelphrey et al., 2002) and auditory (e.g., Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998) sensory modalities.

Consider the results of an eye tracking study in which Pelphrey and colleagues (2002) observed that individuals with autism did not make use of the eyes when judging facial expressions of emotion (see right panels of Figure 1). While repetitive behaviors or language deficits are seen in other disorders (e.g., obsessive-compulsive disorder and specific language impairment, respectively), basic social deficits of this nature are unique to ASD. Onset of the social deficits appears to precede difficulties in other domains (Osterling, Dawson, & Munson, 2002) and may emerge as early as 6 months of age (Maestro et al., 2002).

Defining the Social Brain

Within the past few decades, research has elucidated specific brain circuits that support perception of humans and other species. This *social perception* refers to "the initial stages in the processing of information that culminates in the accurate analysis of the dispositions and intentions of other individuals" (Allison, Puce, & McCarthy, 2000). Basic social perception is a critical building block for more sophisticated social behaviors, such as thinking about the motives and emotions of others. Brothers (1990) first suggested the notion of a **social brain**, a set of interconnected neuroanatomical structures that process social information, enabling the recognition of other individuals and the evaluation their mental states (e.g., intentions, dispositions, desires, and beliefs).

The social brain is hypothesized to consist of the amygdala, the orbital frontal cortex (OFC), fusiform gyrus (FG), and the posterior superior temporal sulcus (STS) region, among other structures. Though all areas work in coordination to support social processing, each appears to serve a distinct role. The amygdala helps us recognize the emotional states of others (e.g., Morris et al., 1996) and also to experience and regulate our own emotions (e.g., LeDoux, 1992). The OFC supports the "reward" feelings we have when we are around other people (e.g., Rolls, 2000). The FG, located at the bottom of the surface of the temporal lobes detects faces and supports face recognition (e.g., Puce, Allison, Asgari, Gore, & McCarthy, 1996). The posterior STS region recognizes the biological motion, including eye, hand and other body movements, and helps to interpret and predict the actions and intentions of others (e.g., Pelphrey, Morris, Michelich, Allison, & McCarthy, 2005).

Current Understanding of Social Perception in ASD

The social brain is of great research interest because the social difficulties characteristic of ASD are thought to relate closely to the functioning of this brain network. Functional magnetic resonance imaging (fMRI) and event-related potentials (ERP) are complementary brain imaging methods used to study activity in the brain across the lifespan. Each method measures a distinct facet of brain activity and contributes unique information to our understanding of brain function.

The human brain has specialized functions to help guide our social interactions. [Image: Allan Ajifo, https://goo.gl/jv4iXf, CC BY 2.0, https://goo.gl/BRvSA7]

FMRI uses powerful magnets to measure the levels of oxygen within the brain, which vary according to changes in neural activity. As the neurons in specific brain regions "work harder", they require more oxygen. FMRI detects the brain regions that exhibit a relative increase in blood flow (and oxygen levels) while people listen to or view social stimuli in the MRI scanner. The areas of the brain most crucial for different social processes are thus identified, with spatial information being accurate to the millimeter.

In contrast, ERP provides direct measurements of the firing of groups of neurons in the cortex. Non-invasive sensors on the scalp record the small electrical currents created by this neuronal activity while the subject views stimuli or listens to specific kinds of information. While fMRI provides information about *where* brain activity occurs, ERP specifies *when* by detailing the timing of processing at the millisecond pace at which it unfolds.

ERP and fMRI are complementary, with fMRI providing excellent *spatial resolution* and ERP offering outstanding *temporal resolution*. Together, this information is critical to understanding the nature of social perception in ASD. To date, the most thoroughly investigated areas of the social brain in ASD are the superior temporal sulcus (STS), which underlies the perception and interpretation of biological motion, and the fusiform gyrus (FG), which supports face perception. Heightened sensitivity to biological motion (for humans, motion such as walking) serves an essential role in the development of humans and other highly social species. Emerging in the first days of life, the ability to detect biological motion helps to orient vulnerable young to critical sources of sustenance, support, and learning, and develops independent of visual experience with biological motion (e.g., Simion, Regolin, & Bulf, 2008).

This inborn "life detector" serves as a foundation for the subsequent development of more complex social behaviors (Johnson, 2006).

From an evolutionary standpoint, it was incredibly important for our survival to maintain social relationships. Therefore, it makes sense that we would be able to recognize faces within the first few days of our infancy. [Image: donnierayjones, https://goo.gl/obrI2x, CC BY 2.0, https://goo.gl/v4Y0Zv]

From very early in life, children with ASD display reduced sensitivity to biological motion (Klin, Lin, Gorrindo, Ramsay, & Jones, 2009). Individuals with ASD have reduced activity in the STS during biological motion perception. Similarly, people at increased genetic risk for ASD but who do not develop symptoms of the disorder (i.e. unaffected siblings of individuals with ASD) show increased activity in this region, which is hypothesized to be a compensatory mechanism to offset genetic vulnerability (Kaiser et al., 2010).

In typical development, preferential attention to faces and the ability to recognize individual faces emerge in the first days of life (e.g., Goren, Sarty, & Wu, 1975). The special way in which the brain responds to faces usually emerges by three months of age (e.g., de Haan, Johnson, & Halit, 2003) and continues throughout the lifespan (e.g., Bentin et al., 1996). Children with ASD, however, tend to show decreased attention to human faces by six to 12 months (Osterling & Dawson, 1994). Children with ASD also show reduced activity in the FG when viewing faces (e.g., Schultz et al., 2000). Slowed processing of faces (McPartland, Dawson, Webb, Panagiotides, & Carver, 2004) is a characteristic of people with ASD that is shared by parents of children with ASD (Dawson, Webb, & McPartland, 2005) and infants at increased risk for developing ASD because of having a sibling with ASD (McCleery, Akshoomoff, Dobkins, & Carver, 2009). Behavioral and attentional differences in face perception and recognition are evident in children and adults with ASD as well (e.g., Hobson, 1986).

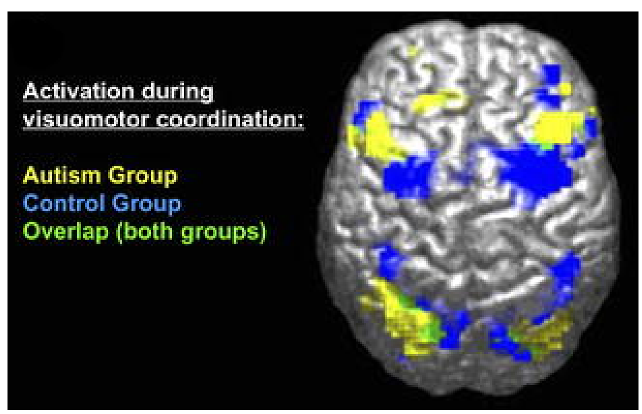
Exploring Diversity in ASD

Because of the limited quality of the behavioral methods used to diagnose ASD and current clinical diagnostic practice, which permits similar diagnoses despite distinct symptom profiles

(McPartland, Webb, Keehn, & Dawson, 2011), it is possible that the group of children currently referred to as having ASD may actually represent different syndromes with distinct causes. Examination of the social brain may well reveal diagnostically meaningful subgroups of children with ASD. Measurements of the "where" and "when" of brain activity during social processing tasks provide reliable sources of the detailed information needed to profile children with ASD with greater accuracy. These profiles, in turn, may help to inform treatment of ASD by helping us to match specific treatments to specific profiles.

The integration of imaging methods is critical for this endeavor. Using face perception as an example, the combination of fMRI and ERP could identify who, of those individuals with ASD, shows anomalies in the FG and then determine the stage of information processing at which these impairments occur. Because different processing stages often reflect discrete cognitive processes, this level of understanding could encourage treatments that address specific processing deficits at the neural level.

For example, differences observed in the early processing stages might reflect problems with low-level visual perception, while later differences would indicate problems with higher-order processes, such as emotion recognition. These same principles can be applied to the broader network of social brain regions and, combined with measures of behavioral functioning, could offer a comprehensive profile of brain-behavior performance for a given individual. A fundamental goal for this kind of subgroup approach is to improve the ability to tailor treatments to the individual.



Trying to diagnose the precise autism disorder can be difficult; many cases share similar symptoms. However, burgeoning technology, like the fMRI, allows clinicians a glimpse into the patient's brain and thus a better understanding of his or her disorder. [Image: Ralph-Axel Müller, https://goo.gl/WwxCV1, CC BY 2.5, https://goo.gl/0QtWcf]

Another objective is to improve the power of other scientific tools. Most studies of individuals with ASD compare groups of individuals, for example, individuals on with ASD compared to typically developing peers. However, studies have also attempted to compare children across the autism spectrum by group according to differential diagnosis (e.g., Asperger's disorder versus autistic disorder), or by other behavioral or cognitive characteristics (e.g., cognitively able versus intellectually disabled or anxious versus non-anxious). Yet, the power of a scientific study to detect these kinds of significant, meaningful, individual differences is only as strong as the accuracy of the factor used to define the compared groups.

The identification of distinct subgroups within the autism spectrum according to information about the brain would allow for a more accurate and detailed exposition of the individual differences seen in those with ASD. This is especially critical for the success of investigations into the genetic basis of ASD. As mentioned before, the genes discovered thus far account for only a small portion of ASD cases. If meaningful, quantitative distinctions in individuals with ASD are identified; a more focused examination into the genetic causes specific to each subgroup could then be pursued. Moreover, distinct findings from neuroimaging, or *biomarkers*, can help guide genetic research. **Endophenotypes**, or characteristics that are not immediately available to observation but that reflect an underlying genetic liability for disease, expose the most basic components of a complex psychiatric disorder and are more stable across the lifespan than observable behavior (Gottesman & Shields, 1973). By describing the key characteristics of ASD in these objective ways, neuroimaging research will facilitate identification of genetic contributions to ASD.

Atypical Brain Development Before the Emergence of Atypical Behavior

Because autism is a developmental disorder, it is particularly important to diagnose and treat ASD early in life. Early deficits in attention to biological motion, for instance, derail subsequent experiences in attending to higher level social information, thereby driving development toward more severe dysfunction and stimulating deficits in additional domains of functioning, such as language development. The lack of reliable predictors of the condition during the first year of life has been a major impediment to the effective treatment of ASD. Without early predictors, and in the absence of a firm diagnosis until behavioral symptoms emerge, treatment is often delayed for two or more years, eclipsing a crucial period in which intervention may be particularly successful in ameliorating some of the social and communicative impairments seen in ASD.

In response to the great need for sensitive (able to identify subtle cases) and specific (able to distinguish autism from other disorders) early indicators of ASD, such as biomarkers, many research groups from around the world have been studying patterns of infant development using prospective longitudinal studies of infant siblings of children with ASD and a comparison group of infant siblings without familial risks. Such designs gather longitudinal information about developmental trajectories across the first three years of life for both groups followed by clinical diagnosis at approximately 36 months.

These studies are problematic in that many of the social features of autism do not emerge in

If autism is diagnosed early enough, treatments have developed to the point that children with ASD can learn and grow to have more intensive social interactions. [Image: hepingting, https://goo.gl/TIoAcY, CC BY-SA 2.0, https://goo.gl/rxiUsF]

typical development until after 12 months of age, and it is not certain that these symptoms will manifest during the limited periods of observation involved in clinical evaluations or in pediatricians' offices. Moreover, across development, but especially during infancy, behavior is widely variable and often unreliable, and at present, behavioral observation is the only means to detect symptoms of ASD and to confirm a diagnosis. This is quite problematic because, even highly sophisticated behavioral methods, such as eye tracking (see Figure 1), do not necessarily reveal reliable differences in infants with ASD (Ozonoff et al., 2010). However, measuring the brain activity associated with social perception can detect differences that do not appear in behavior until much later. The identification of biomarkers utilizing the imaging methods we have described offers promise for earlier detection of atypical social development.

ERP measures of brain response predict subsequent development of autism in infants as young as six months old who showed normal patterns of visual fixation (as measured by eye tracking) (Elsabbagh et al., 2012). This suggests the great promise of brain imaging for earlier recognition of ASD. With earlier detection, treatments could move from addressing existing symptoms to preventing their emergence by altering the course of abnormal brain development and steering it toward normality.

Hope for Improved Outcomes

The brain imaging research described above offers hope for the future of ASD treatment. Many of the functions of the social brain demonstrate significant *plasticity*, meaning that their functioning can be affected by experience over time. In contrast to theories that suggest difficulty processing complex information or communicating across large expanses of cortex (Minshew & Williams, 2007), this malleability of the social brain is a positive prognosticator for the development of treatment. The brains of people with ASD are not wired to process optimally social information. But this does not mean that these systems are irretrievably *broken*. Given the observed plasticity of the social brain, remediation of these difficulties may be possible with appropriate and timely intervention.

Outside Resources

Web: American Psychiatric Association's website for the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders
http://www.dsm5.org

Web: Autism Science Foundation - organization supporting autism research by providing funding and other assistance to scientists and organizations conducting, facilitating, publicizing and disseminating autism research. The organization also provides information about autism to the general public and serves to increase awareness of autism spectrum disorders and the needs of individuals and families affected by autism.
http://www.autismsciencefoundation.org/

Web: Autism Speaks - Autism science and advocacy organization
http://www.autismspeaks.org/

Discussion Questions

1. How can neuroimaging inform our understanding of the causes of autism?
2. What are the ways in which neuroimaging, including fMRI and ERP, may benefit efforts to diagnosis and treat autism?
3. How can an understanding of the social brain help us to understand ASD?
4. What are the core symptoms of ASD, and why is the social brain of particular interest?
5. What are some of the components of the social brain, and what functions do they serve?

Vocabulary

Endophenotypes
A characteristic that reflects a genetic liability for disease and a more basic component of a complex clinical presentation. Endophenotypes are less developmentally malleable than overt behavior.

Event-related potentials (ERP)
Measures the firing of groups of neurons in the cortex. As a person views or listens to specific types of information, neuronal activity creates small electrical currents that can be recorded from non-invasive sensors placed on the scalp. ERP provides excellent information about the timing of processing, clarifying brain activity at the millisecond pace at which it unfolds.

Functional magnetic resonance imaging (fMRI)
Entails the use of powerful magnets to measure the levels of oxygen within the brain that vary with changes in neural activity. That is, as the neurons in specific brain regions "work harder" when performing a specific task, they require more oxygen. By having people listen to or view social percepts in an MRI scanner, fMRI specifies the brain regions that evidence a relative increase in blood flow. In this way, fMRI provides excellent spatial information, pinpointing with millimeter accuracy, the brain regions most critical for different social processes.

Social brain
The set of neuroanatomical structures that allows us to understand the actions and intentions of other people.

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Science, 4*(7), 267–278.

Bentin, S., Allison, T., Puce, A., Perez, E., et al. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*(6), 551–565.

Brothers, L. (1990). The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience, 1*, 27–51.

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism & Developmental Disorders, 28*(6), 479–485.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology, 27*(3), 403–424.

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., et al. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology, 22*(4), 338–342.

Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: Developmental disconnection syndromes. *Current Opinion in Neurobiology, 17*(1), 103–111.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56*(4), 544–549.

Gottesman I. I., & Shields, J. (1973) Genetic theorizing and schizophrenia. *British Journal of Psychiatry, 122*, 15–30.

Hobson, R. (1986). The autistic child's appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry, 27*(3), 321–342.

Johnson, M. H. (2006). Biological motion: A perceptual life detector? *Current Biology, 16*(10), R376–377.

Kaiser, M. D., Hudac, C. M., Shultz, S., Lee, S. M., Cheung, C., Berken, A. M., et al. (2010). Neural signatures of autism. *Proceedings of the National Academy of Sciences of the United States of America, 107*(49), 21223–21228.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature, 459*(7244), 257–261.

Maestro, S., Muratori, F., Cavallaro, M. C., Pei, F., Stern, D., Golse, B., et al. (2002). Attentional skills during the first 6 months of age in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(10), 1239–1245.

McCleery, J. P., Akshoomoff, N., Dobkins, K. R., & Carver, L. J. (2009). Atypical face versus object processing and hemispheric asymmetries in 10-month-old infants at risk for autism. *Biological Psychiatry, 66*(10), 950–957.

McPartland, J. C., Dawson, G., Webb, S. J., Panagiotides, H., & Carver, L. J. (2004). Event-related brain potentials reveal anomalies in temporal processing of faces in autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 45*(7), 1235–1245.

McPartland, J. C., Webb, S. J., Keehn, B., & Dawson, G. (2011). Patterns of visual attention to faces and objects in autism spectrum disorder. *Journal of Autism and Develop Disorders, 41* (2), 148–157.

Minshew, N. J., & Williams, D. L. (2007). The new neurobiology of autism: Cortex, connectivity, and neuronal organization. *Archives of Neurology, 64*(7), 945–950.

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders, 24*, 247-257.

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development & Psychopathology, 14* (2), 239–251.

Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., et al. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(3), 256–266.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism & Developmental Disorders, 32*(4), 249–261.

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., et al. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archives of General Psychiatry, 57*(4), 331–340.

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences, 105*(2), 809–813.

de Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: A review. *International Journal of Psychophysiology, 51*(1), 45–58.

39
Psychopharmacology

Susan Barron

Psychopharmacology is the study of how drugs affect behavior. If a drug changes your perception, or the way you feel or think, the drug exerts effects on your brain and nervous system. We call drugs that change the way you think or feel psychoactive or psychotropic drugs, and almost everyone has used a psychoactive drug at some point (yes, caffeine counts). Understanding some of the basics about psychopharmacology can help us better understand a wide range of things that interest psychologists and others. For example, the pharmacological treatment of certain neurodegenerative diseases such as Parkinson's disease tells us something about the disease itself. The pharmacological treatments used to treat psychiatric conditions such as schizophrenia or depression have undergone amazing development since the 1950s, and the drugs used to treat these disorders tell us something about what is happening in the brain of individuals with these conditions. Finally, understanding something about the actions of drugs of abuse and their routes of administration can help us understand why some psychoactive drugs are so addictive. In this module, we will provide an overview of some of these topics as well as discuss some current controversial areas in the field of psychopharmacology.

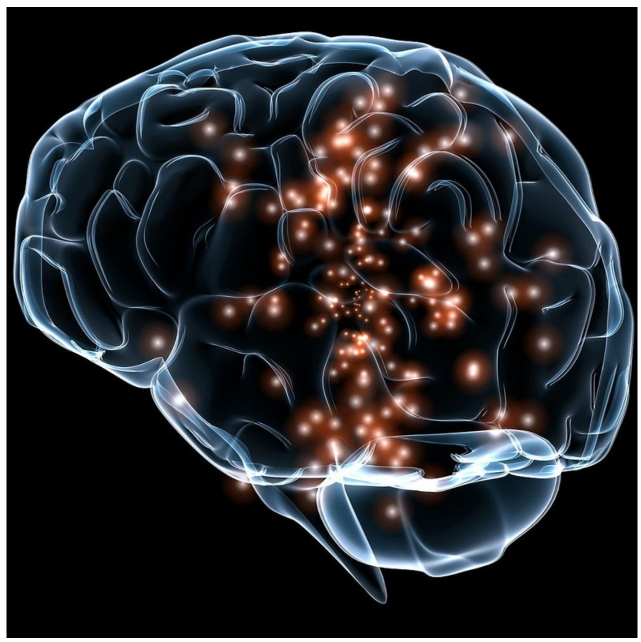
Learning Objectives

- How do the majority of psychoactive drugs work in the brain?
- How does the route of administration affect how rewarding a drug might be?
- Why is grapefruit dangerous to consume with many psychotropic medications?
- Why might individualized drug doses based on genetic screening be helpful for treating conditions like depression?
- Why is there controversy regarding pharmacotherapy for children, adolescents, and the elderly?

Introduction

Psychopharmacology, the study of how drugs affect the brain and behavior, is a relatively new science, although people have probably been taking drugs to change how they feel from early in human history (consider the of eating fermented fruit, ancient beer recipes, chewing on the leaves of the cocaine plant for stimulant properties as just some examples). The word *psychopharmacology* itself tells us that this is a field that bridges our understanding of behavior (and brain) and pharmacology, and the range of topics included within this field is extremely broad.

Virtually any drug that changes the way you feel does this by altering how neurons communicate with each other. Neurons (86 billion in your nervous system) communicate with each other by releasing a chemical (neurotransmitter) across a tiny space between two neurons (the synapse). When the neurotransmitter crosses the synapse, it binds to a postsynaptic receptor (protein) on the receiving neuron and the message may then be transmitted onward. Obviously, neurotransmission is far more complicated than this – links at the end of this module can provide some useful background if you want more detail – but the first step is understanding that virtually all psychoactive drugs interfere with or alter how neurons communicate with each other.



Drugs that alter our feelings and behavior do so by affecting the communication between neurons in the brain. [Image: https://goo.gl/oQCafL, CC0 Public Domain, https://goo.gl/m25gce]

There are many neurotransmitters. Some of the most important in terms of psychopharmacological treatment and drugs of abuse are outlined in Table 1. The neurons that release these neurotransmitters, for the most part, are localized within specific circuits of the brain that mediate these behaviors. Psychoactive drugs can either increase activity at the synapse (these are called agonists) or reduce activity at the synapse (antagonists). Different drugs do this by different mechanisms, and some examples of agonists and antagonists are presented in Table 2. For each example, the drug's trade name, which is the name of the drug provided by the drug company, and generic name (in parentheses) are provided.

Neurotransmitter	Abbreviation	Behaviors or Diseases Related to These Neurotransmitter
Acetylcholine	ACh	Learning and memory; Alzheimer's disease' muscle movement in the peripheral nervous system
Dopamine	DA	Reward circuits; Motor circuits involved in Parkinson's disease; Schizophrenia
Norepinephrine	NE	Arousal; Depression
Serotonin	5HT	Depression; Aggression; Schizophrenia
Glutamate	GLU	Learning; Major excitatory neurotransmitter in the brain
GABA	GABA	Anxiety disorders; Epilepsy; Major inhibitory neurotransmitter in the brain
Endogenous Opiods	Endorphins, Enkephalins	Pain; Analgesia; Reward

Table 1

A very useful link at the end of this module shows the various steps involved in neurotransmission and some ways drugs can alter this.

Table 2 provides examples of drugs and their primary mechanism of action, but it is very important to realize that drugs also have effects on other neurotransmitters. This contributes to the kinds of side effects that are observed when someone takes a particular drug. The reality is that no drugs currently available work only exactly where we would like in the brain or only on a specific neurotransmitter. In many cases, individuals are sometimes prescribed one psychotropic drug but then may also have to take additional drugs to reduce the side effects caused by the initial drug. Sometimes individuals stop taking medication because the side effects can be so profound.

Pharmacokinetics: What Is It – Why Is It Important?

While this section may sound more like pharmacology, it is important to realize how important pharmacokinetics can be when considering psychoactive drugs. Pharmacokinetics refers to how the body handles a drug that we take. As mentioned earlier, psychoactive drugs exert

Drug	Mechanism	Use	Agonist/Antagonist
L-dopa	Increase synthesis of DA	Parkinson's disease	Agonist for DA
Adderall (mixed salts amphetamine)	Increase release of DA, NE	ADHD	Agonist for DA, NE
Ritalin (methylphenidate)	Blocks removal of DA, NE, and lesser (5HT) from synapse	ADHD	Agonist for DA, NE mostly
Aricept (donepezil)	Blocks removal of ACh from synapse	Alzheimer's disease	Agonist for ACh
Prozac (fluoxetine)	Blocks removal of 5HT from synapse	Depression, obsessive compulsive disorder	Agonist 5HT
Seroquel (quetiapine)	Blocks DA and 5HT receptors	Schizophrenia, bipolar disorder	Antagonist for DA, 5HT
Revia (naltrexone)	Blocks opiod post-synaptic receptors	Alcoholism, opiod addiction	Antagonist (for opioids)

Table 2

their effects on behavior by altering neuronal communication in the brain, and the majority of drugs reach the brain by traveling in the blood. The acronym ADME is often used with A standing for absorption (how the drug gets into the blood), Distribution (how the drug gets to the organ of interest – in this module, that is the brain), Metabolism (how the drug is broken down so it no longer exerts its psychoactive effects), and Excretion (how the drug leaves the body). We will talk about a couple of these to show their importance for considering psychoactive drugs.

A drug delivered by IV reaches the brain more quickly than if the drug is taken orally. While rapid delivery has advantages, there are also risks involved with IV administration. [Image: Calleamanecer, https://goo.gl/OX6Yj5, CC BY-SA 3.0, https://goo.gl/eLCn2O]

Drug Administration

There are many ways to take drugs, and these routes of drug administration can have a significant impact on how quickly that drug reaches brain. The most common route of administration is oral administration, which is relatively slow and – perhaps surprisingly – often the most variable and complex route of administration. Drugs enter the stomach and then get absorbed by the blood supply and capillaries that line the small intestine. The rate of absorption can be affected by a variety of factors including the quantity and the type of food in the stomach (e.g., fats vs. proteins). This is why the medicine label for some drugs (like antibiotics) may specifically state foods that you should or should NOT consume within an hour of taking the drug because they can affect the rate of absorption. Two of the most rapid routes of administration include inhalation (i.e., smoking or gaseous anesthesia) and intravenous (IV) in which the drug is injected directly into the vein and hence the blood supply. Both of these routes of administration can get the drug to brain in less than 10 seconds. IV administration also has the distinction of being the most dangerous because if there is an adverse drug reaction, there is very little time to administer any antidote, as in the case of an IV heroin overdose.

Why might how quickly a drug gets to the brain be important? If a drug activates the reward circuits in the brain AND it reaches the brain very quickly, the drug has a high risk for abuse and addiction. Psychostimulants like amphetamine or cocaine are examples of drugs that have high risk for abuse because they are agonists at DA neurons involved in reward AND because these drugs exist in forms that can be either smoked or injected intravenously. Some argue that cigarette smoking is one of the hardest addictions to quit, and although part of the reason for this may be that smoking gets the nicotine into the brain very quickly (and indirectly acts on DA neurons), it is a more complicated story. For drugs that reach the brain very quickly, not only is the drug very addictive, but so are the cues associated with the drug (see Rohsenow, Niaura, Childress, Abrams, & Monti, 1990). For a crack user, this could be the pipe that they use to smoke the drug. For a cigarette smoker, however, it could be something as normal as finishing dinner or waking up in the morning (if that is when the smoker usually has a cigarette). For both the crack user and the cigarette smoker, the cues associated with the drug may actually cause craving that is alleviated by (you guessed it) – lighting a cigarette or using crack (i.e., relapse). This is one of the reasons individuals that enroll in drug treatment programs, especially out-of-town programs, are at significant risk of relapse if they later find themselves in proximity to old haunts, friends, etc. But this is much *more* difficult for a cigarette smoker. How can someone avoid eating? Or avoid waking up in the morning, etc. These examples help you begin to understand how important the route of administration can be for psychoactive drugs.

Drug Metabolism

Metabolism involves the breakdown of psychoactive drugs, and this occurs primarily in the liver. The liver produces enzymes (proteins that speed up a chemical reaction), and these enzymes help catalyze a chemical reaction that breaks down psychoactive drugs. Enzymes exist in "families," and many psychoactive drugs are broken down by the same family of enzymes, the cytochrome P450 superfamily. There is not a unique enzyme for each drug; rather, certain enzymes can break down a wide variety of drugs. Tolerance to the effects of many drugs can occur with repeated exposure; that is, the drug produces less of an effect over time, so more of the drug is needed to get the same effect. This is particularly true for sedative drugs like alcohol or opiate-based painkillers. *Metabolic tolerance* is one kind of tolerance and it takes place in the liver. Some drugs (like alcohol) cause **enzyme induction** – an increase in the enzymes produced by the liver. For example, chronic drinking results in alcohol being broken down more quickly, so the alcoholic needs to drink more to get the same effect – of course, until so much alcohol is consumed that it damages the liver (alcohol can cause fatty liver or cirrhosis).

Recent Issues Related to Psychotropic Drugs and Metabolism

Grapefruit Juice and Metabolism

Grapefruit can interfere with enzymes in the liver that help the body to process certain drugs. [Image: CC0 Public Domain, https://goo.gl/m25gce]

Certain types of food in the stomach can alter the rate of drug absorption, and other foods can also alter the rate of drug metabolism. The most well known is grapefruit juice. Grapefruit juice suppresses cytochrome P450 enzymes in the liver, and these liver enzymes normally break down a large variety of drugs (including some of the psychotropic drugs). If the enzymes are suppressed, drug levels can build up to potentially toxic levels. In this case, the effects can persist for extended periods of time after the consumption of grapefruit juice. As of 2013, there are at least 85 drugs shown to adversely interact with grapefruit juice (Bailey, Dresser, & Arnold, 2013). Some psychotropic drugs that are likely to interact with grapefruit juice include carbamazepine (Tegretol), prescribed for bipolar disorder; diazepam (Valium), used to treat anxiety, alcohol withdrawal, and muscle spasms; and fluvoxamine (Luvox), used to treat obsessive compulsive disorder and depression. A link at the end of this module gives the latest list of drugs reported to have this unusual interaction.

Individualized Therapy, Metabolic Differences, and Potential Prescribing Approaches for the Future

Mental illnesses contribute to more disability in western countries than all other illnesses including cancer and heart disease. Depression alone is predicted to be the second largest contributor to disease burden by 2020 (World Health Organization, 2004). The numbers of people affected by mental health issues are pretty astonishing, with estimates that 25% of adults experience a mental health issue in any given year, and this affects not only the individual but their friends and family. One in 17 adults experiences a serious mental illness (Kessler, Chiu, Demler, & Walters, 2005). Newer antidepressants are probably the most frequently prescribed drugs for treating mental health issues, although there is no “magic bullet” for treating depression or other conditions. Pharmacotherapy with psychological therapy may be the most beneficial treatment approach for many psychiatric conditions, but there are still many unanswered questions. For example, why does one antidepressant help one individual yet have no effect for another? Antidepressants can take 4 to 6 weeks to start improving depressive symptoms, and we don’t really understand why. Many people do not respond to the first antidepressant prescribed and may have to try different drugs before finding something that works for them. Other people just do not improve with antidepressants (Ioannidis, 2008). As we better understand why individuals differ, the easier and more rapidly we will be able to help people in distress.

One area that has received interest recently has to do with an individualized treatment approach. We now know that there are genetic differences in some of the cytochrome P450 enzymes and their ability to break down drugs. The general population falls into the following 4 categories: 1) *ultra-extensive metabolizers* break down certain drugs (like some of the current antidepressants) very, very quickly, 2) *extensive metabolizers* are also able to break down drugs fairly quickly, 3) *intermediate metabolizers* break down drugs more slowly than either of the two above groups, and finally 4) *poor metabolizers* break down drugs much more slowly than all of the other groups. Now consider someone receiving a prescription for an antidepressant – what would the consequences be if they were either an ultra-extensive metabolizer or a poor metabolizer? The ultra-extensive metabolizer would be given antidepressants and told it will probably take 4 to 6 weeks to begin working (this is true), but they metabolize the medication so quickly that it will never be effective for them. In contrast, the poor metabolizer given the same daily dose of the same antidepressant may build up such high levels in their blood (because they are not breaking the drug down), that they will have a wide range of side effects and feel really badly – also not a positive outcome. What if – instead – prior to prescribing an antidepressant, the doctor could take a blood sample and determine which type of metabolizer a patient actually was? They could then make a much more informed decision about the best dose to prescribe. There are new genetic tests now available to better individualize treatment in just this way. A blood sample can determine (at least for some drugs) which category an individual fits into, but we need data to determine if this actually is effective for treating depression or other mental illnesses (Zhou, 2009). Currently, this genetic test is expensive and not many health insurance plans cover this screen, but this may be an important component in the future of psychopharmacology.

Other Controversial Issues

Juveniles and Psychopharmacology

A recent Centers for Disease Control (CDC) report has suggested that as many as 1 in 5 children between the ages of 5 and 17 may have some type of mental disorder (e.g., ADHD, autism, anxiety, depression) (CDC, 2013). The incidence of bipolar disorder in children and adolescents has also increased 40 times in the past decade (Moreno, Laje, Blanco, Jiang, Schmidt, & Olfson, 2007), and it is now estimated that 1 in 88 children have been diagnosed with an autism spectrum disorder (CDC, 2011). Why has there been such an increase in these numbers? There is no single answer to this important question. Some believe that greater public awareness has contributed to increased teacher and parent referrals. Others argue that the increase stems from changes in criterion currently used for diagnosing. Still others suggest environmental factors, either prenatally or postnatally, have contributed to this upsurge.

We do not have an answer, but the question does bring up an additional controversy related to how we should treat this population of children and adolescents. Many psychotropic drugs used for treating psychiatric disorders have been tested in adults, but few have been tested for safety

There are concerns about both the safety and efficacy of drugs like Prozac for children and teens. [Image: zaza_bj, CC BY-NC-SA 2.0, https://goo.gl/Toc0ZF] or efficacy with children or adolescents. The most well-established psychotropics prescribed for children and adolescents are the psychostimulant drugs used for treating attention deficit hyperactivity disorder (ADHD), and there are clinical data on how effective these drugs are. However, we know far less about the safety and efficacy in young populations of the drugs typically prescribed for treating anxiety, depression, or other psychiatric disorders. The young brain continues to mature until probably well after age 20, so some scientists are concerned that drugs that alter neuronal activity in the developing brain could have significant consequences. There is an obvious need for clinical trials in children and adolescents to test the safety and effectiveness of many of these drugs, which also brings up a variety of ethical questions about who decides what children and adolescents will participate in these clinical trials, who can give consent, who receives reimbursements, etc.

The Elderly and Psychopharmacology

Another population that has not typically been included in clinical trials to determine the safety or effectiveness of psychotropic drugs is the elderly. Currently, there is very little high-quality evidence to guide prescribing for older people – clinical trials often exclude people with multiple comorbidities (other diseases, conditions, etc.), which are typical for elderly populations (see Hilmer and Gnjidict, 2008; Pollock, Forsyth, & Bies, 2008). This is a serious issue because the elderly consume a disproportionate number of the prescription meds prescribed. The term polypharmacy refers to the use of multiple drugs, which is very common

Clinical Trial Subjects	Aged Patients Who Receive Drug Therapies
One drug	Drug of interest and medications
Single does	Chronic administration
No disease	Multiple diseases
No alcohol, tobacco, OTC* drugs, nutraceuticals	OTC* drugs, nutraceuticals, alcohol, tobacco, and other
20-40 years (vs 60-75 years)	65-100+ years
Caucaisians	Caucasians and minorities
Selection bias	All comers/socioeconomic basis

*OTC = Over the counter

Table 3. Characteristics if clinical trial subjects vs. actual patients. (Reprinted by permission from Schwartz & Abernethy, 2008.)

in elderly populations in the United States. As our population ages, some estimate that the proportion of people 65 or older will reach 20% of the U.S. population by 2030, with this group consuming 40% of the prescribed medications. As shown in Table 3 (from Schwartz and Abernethy, 2008), it is quite clear why the typical clinical trial that looks at the safety and effectiveness of psychotropic drugs can be problematic if we try to interpret these results for an elderly population.

Metabolism of drugs is often slowed considerably for elderly populations, so less drug can produce the same effect (or all too often, too much drug can result in a variety of side effects). One of the greatest risk factors for elderly populations is falling (and breaking bones), which can happen if the elderly person gets dizzy from too much of a drug. There is also evidence that psychotropic medications can reduce bone density (thus worsening the consequences if someone falls) (Brown & Mezuk, 2012). Although we are gaining an awareness about some of the issues facing pharmacotherapy in older populations, this is a very complex area with many medical and ethical questions.

This module provided an introduction of some of the important areas in the field of psychopharmacology. It should be apparent that this module just touched on a number of topics included in this field. It should also be apparent that understanding more about psychopharmacology is important to anyone interested in understanding behavior and that our understanding of issues in this field has important implications for society.

Outside Resources

Video: Neurotransmission
http://www.youtube.com/watch?v=FR4S1BqdFG4

Web: Description of how some drugs work and the brain areas involved - 1
http://www.drugabuse.gov/news-events/nida-notes/2007/10/impacts-drugs-neurotransmission

Web: Description of how some drugs work and the brain areas involved - 2
http://learn.genetics.utah.edu/content/addiction/mouse/

Web: Information about how neurons communicate and the reward pathways
http://learn.genetics.utah.edu/content/addiction/rewardbehavior/

Web: National Institute of Alcohol Abuse and Alcoholism
http://www.niaaa.nih.gov/

Web: National Institute of Drug Abuse
http://www.drugabuse.gov/

Web: National Institute of Mental Health
http://www.nimh.nih.gov/index.shtml

Web: Neurotransmission
https://science.education.nih.gov/supplements/nih2/Addiction/activities/lesson2_neurotransmission.html

Web: Report of the Working Group on Psychotropic Medications for Children and Adolescents: Psychopharmacological, Psychosocial, and Combined Interventions for Childhood Disorders: Evidence Base, Contextual Factors, and Future Directions (2008):
http://www.apa.org/pi/families/resources/child-medications.pdf

Web: Ways drugs can alter neurotransmission
http://thebrain.mcgill.ca/flash/d/d_03/d_03_m/d_03_m_par/d_03_m_par.html

Discussion Questions

1. What are some of the issues surrounding prescribing medications for children and

adolescents? How might this be improved?

2. What are some of the factors that can affect relapse to an addictive drug?
3. How might prescribing medications for depression be improved in the future to increase the likelihood that a drug would work and minimize side effects?

Vocabulary

Agonists
A drug that increases or enhances a neurotransmitter's effect.

Antagonist
A drug that blocks a neurotransmitter's effect.

Enzyme
A protein produced by a living organism that allows or helps a chemical reaction to occur.

Enzyme induction
Process through which a drug can enhance the production of an enzyme.

Metabolism
Breakdown of substances.

Neurotransmitter
A chemical substance produced by a neuron that is used for communication between neurons.

Pharmacokinetics
The action of a drug through the body, including absorption, distribution, metabolism, and excretion.

Polypharmacy
The use of many medications.

Psychoactive drugs
A drug that changes mood or the way someone feels.

Psychotropic drug
A drug that changes mood or emotion, usually used when talking about drugs prescribed for various mental conditions (depression, anxiety, schizophrenia, etc.).

Synapse
The tiny space separating neurons.

References

Bailey D. G., Dresser G., & Arnold J. M. (2013). Grapefruit-medication interactions: forbidden fruit or avoidable consequences? *Canadian Medical Association Journal, 185*, 309–316.

Brown, M. J., & Mezuk, B. (2012). Brains, bones, and aging: psychotropic medications and bone health among older adults. *Current Osteoporosis Reports, 10*, 303–311.

Centers for Disease Control and Prevention (2011) Prevalence of autism spectrum disorders – autism and developmental disabilities monitoring network, 14 sites, United States, 2008. *Morbidity and Mortality Weekly Report 61*(SS03) 1–19.

Centers for Disease Control and Prevention. (2013) Mental health surveillance among children – United States, 2005—2011. *Morbidity and Mortality Weekly Report 62* Suppl, 1-35.

Hilmer, N., & Gnjidict, D. (2008). The effects of polypharmacy in older adults. *Clinical Pharmacology & Therapeutics, 85*, 86–88.

Ioannidis, J. P. A. (2008). Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philosophy, Ethics and Humanities in Medicine, 3*,14.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of twelve-month DSM-IV disorders in the National Comorbidity Survey Replication (NCS-R). *Archives of General Psychiatry, 62*, 617–627.

Moreno, C., Laje, G., Blanco, C., Jiang, H., Schmidt, A. B., & Olfson, M., (2007). National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. *Archives of General Psychiatry, 64*(9), 1032–1039.

Pollock, B. G., Forsyth, C. E., & Bies, R. R. (2008). The critical role of clinical pharmacology in geriatric psychopharmacology. *Clinical Pharmacology & Therapeutics, 85*, 89–93.

Rohsenow, D. J., Niaura, R. S., Childress, A. R., Abrams, D. B., &, Monti, P. M. (1990). Cue reactivity in addictive behaviors: Theoretical and treatment implications. *International Journal of Addiction, 25*, 957–993.

Schwartz, J. B., & Abernethy, D. R. (2008). Aging and medications: Past, present, future. *Clinical Pharmacology & Therapeutics, 85*, 3–10.

World Health Organization. (2004). *Promoting mental health: concepts, emerging evidence, practice* (Summary Report). Geneva, Switzerland: Author. Retrieved from http://www.who.int/mental_health/evidence/en/promoting_mhh.pdf

Zhou, S. F. (2009). Polymorphism of human cytochrome P450 2D6 and its clinical significance: Part II. *Clinical Pharmacokinetics, 48*, 761–804.

Well-Being

40
The Healthy Life

Emily Hooker & Sarah Pressman

Our emotions, thoughts, and behaviors play an important role in our health. Not only do they influence our day-to-day health practices, but they can also influence how our body functions. This module provides an overview of health psychology, which is a field devoted to understanding the connections between psychology and health. Discussed here are examples of topics a health psychologist might study, including stress, psychosocial factors related to health and disease, how to use psychology to improve health, and the role of psychology in medicine.

Learning Objectives

- Describe basic terminology used in the field of health psychology.
- Explain theoretical models of health, as well as the role of psychological stress in the development of disease.
- Describe psychological factors that contribute to resilience and improved health.
- Defend the relevance and importance of psychology to the field of medicine.

What Is Health Psychology?

Today, we face more **chronic disease** than ever before because we are living longer lives while also frequently behaving in unhealthy ways. One example of a chronic disease is coronary heart disease (CHD): It is the number one cause of death worldwide (World Health Organization, 2013). CHD develops slowly over time and typically appears midlife, but related heart problems can persist for years after the original diagnosis or cardiovascular event. In managing illnesses that persist over time (other examples might include cancer, diabetes, and long-term disability) many psychological factors will determine the progression of the ailment. For example, do patients seek help when appropriate? Do they follow doctor recommendations? Do they develop negative psychological symptoms due to lasting illness (e.g., depression)? Also important is that psychological factors can play a significant role in *who* develops these diseases, the prognosis, and the nature of the symptoms related to the illness. Health psychology is a relatively new, interdisciplinary field of study that focuses on these very issues, or more specifically, the role of psychology in maintaining health, as well as preventing and treating illness.

Consideration of how psychological and social factors influence health is especially important today because many of the leading causes of illness in developed countries are often attributed to psychological and behavioral factors. In the case of CHD, discussed above, psychosocial factors, such as excessive stress, smoking, unhealthy eating habits, and some personality traits can also lead to increased risk of disease and worse health outcomes. That being said, many of these factors can be adjusted using psychological techniques. For example, clinical health psychologists can improve health practices like poor dietary choices and smoking, they can teach important stress reduction techniques, and they can help treat psychological disorders tied to poor health. Health psychology considers how the choices we make, the behaviors we engage in, and even the emotions that we feel, can play an important role in our overall health (Cohen & Herbert, 1996; Taylor, 2012).

Health psychologists are helping people to adapt behaviors to avoid disease, reduce stress, and improve overall health. [Image: Adelphi Lab Center, https://goo.gl/N9wXon, CC BY 2.0, https://goo.gl/BRvSA7]

Health psychology relies on the **Biopsychosocial Model of Health**. This model posits that biology, psychology, and social factors are just as important in the development of disease as biological causes (e.g., germs, viruses), which is consistent with the World Health Organization (1946) definition of **health**. This model replaces the older **Biomedical Model of Health**, which primarily considers the physical, or pathogenic, factors contributing to illness. Thanks to advances in medical technology, there is a growing understanding of the physiology underlying the **mind–body connection**, and in particular, the role that different feelings can have on our body's function. Health psychology researchers working in the fields of **psychosomatic medicine** and **psychoneuroimmunology**, for example, are interested in understanding how psychological factors can "get under the skin" and influence our physiology in order to better understand how factors like stress can make us sick.

Stress And Health

You probably know exactly what it's like to feel stress, but what you may not know is that it can objectively influence your health. Answers to questions like, "How stressed do you feel?" or "How overwhelmed do you feel?" can predict your likelihood of developing both minor illnesses as well as serious problems like future heart attack (Cohen, Janicki-Deverts, & Miller, 2007). (Want to measure your own stress level? Check out the links at the end of the module.) To understand how health psychologists study these types of associations, we will describe one famous example of a stress and health study. Imagine that you are a research subject for a moment. After you check into a hotel room as part of the study, the researchers ask you to report your general levels of stress. Not too surprising; however, what happens next is that you receive droplets of *cold virus* into your nose! The researchers intentionally try to make you sick by exposing you to an infectious illness. After they expose you to the virus, the researchers will then evaluate you for several days by asking you questions about your symptoms, monitoring how much mucus you are producing by weighing your used tissues, and taking body fluid samples—all to see if you are objectively ill with a cold. Now, the

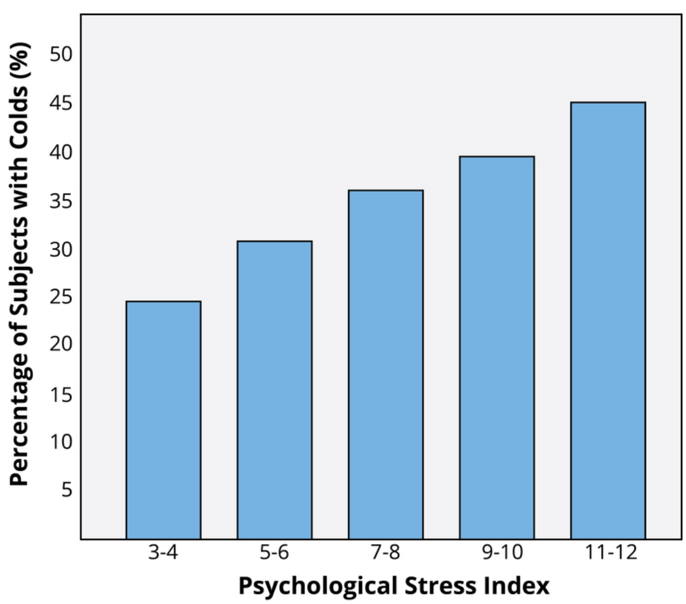


Figure 1: Adapted from Cohen et al. 1991

interesting thing is that not everyone who has drops of cold virus put in their nose develops the illness. Studies like this one find that people who are less stressed and those who are more positive at the beginning of the study are at a decreased risk of developing a cold (Cohen, Tyrrell, & Smith, 1991; Cohen, Alper, Doyle, Treanor, & Turner, 2006) (see Figure 1 for an example).

Importantly, it is not just major life **stressors** (e.g., a family death, a natural disaster) that increase the likelihood of getting sick. Even small **daily hassles** like getting stuck in traffic or fighting with your girlfriend can raise your blood pressure, alter your stress hormones, and even suppress your immune system function (DeLongis, Folkman, & Lazarus, 1988; Twisk, Snel, Kemper, & van Machelen, 1999).

It is clear that stress plays a major role in our mental and physical health, but what exactly is it? The term **stress** was originally derived from the field of mechanics where it is used to describe materials under pressure. The word was first used in a *psychological* manner by researcher Hans Selye. He was examining the effect of an ovarian hormone that he thought caused sickness in a sample of rats. Surprisingly, he noticed that almost any injected hormone produced this same sickness. He smartly realized that it was not the hormone under investigation that was causing these problems, but instead, the aversive experience of being handled and injected by researchers that led to high physiological arousal and, eventually, to health problems like ulcers. Selye (1946) coined the term stressor to label a stimulus that had this effect on the body and developed a model of the stress response called the **General Adaptation Syndrome**. Since then, psychologists have studied stress in a myriad of ways, including stress as negative events (e.g., natural disasters or major life changes like dropping out of school), as chronically difficult situations (e.g., taking care of a loved one with Alzheimer's), as short-term hassles, as a biological fight-or-flight response, and even as clinical illness like post-traumatic stress disorder (PTSD). It continues to be one of the most important and well-studied psychological correlates of illness, because excessive stress causes potentially damaging wear and tear on the body and can influence almost any imaginable disease process.

Protecting Our Health

An important question that health psychologists ask is: What keeps us protected from disease and alive longer? When considering this issue of **resilience** (Rutter, 1985), five factors are often studied in terms of their ability to protect (or sometimes harm) health. They are:

1. Coping

2. Control and Self-Efficacy
3. Social Relationships
4. Dispositions and Emotions
5. Stress Management

Coping Strategies

How individuals cope with the stressors they face can have a significant impact on health. Coping is often classified into two categories: problem-focused coping or emotion-focused coping (Carver, Scheier, & Weintraub, 1989). Problem-focused coping is thought of as actively addressing the event that is causing stress in an effort to solve the issue at hand. For example, say you have an important exam coming up next week. A problem-focused strategy might be to spend additional time over the weekend studying to make sure you understand all of the material. Emotion-focused coping, on the other hand, regulates the emotions that come with stress. In the above examination example, this might mean watching a funny movie to take your mind off the anxiety you are feeling. In the short term, emotion-focused coping might reduce feelings of stress, but problem-focused coping seems to have the greatest impact on mental wellness (Billings & Moos, 1981; Herman-Stabl, Stemmler, & Petersen, 1995). That being said, when events are uncontrollable (e.g., the death of a loved one), emotion-focused coping directed at managing your feelings, at first, might be the better strategy. Therefore, it is always important to consider the match of the stressor to the coping strategy when evaluating its plausible benefits.

Feeling a sense of control in one's life is important. Something as simple as having control over the care of a houseplant has been shown to improve health and longevity. [Image: JJ Harrison, https://goo.gl/82FsdV, CC BY-SA 2.5, https://goo.gl/SRAIwa]

Control and Self-Efficacy

Another factor tied to better health outcomes and an improved ability to cope with stress is having the belief that you have control over a situation. For example, in one study where participants were forced to listen to unpleasant (stressful) noise, those who were led to believe that they had control over the noise performed much better on proofreading tasks afterwards (Glass & Singer, 1972). In other words, even though participants *did not* have actual control over the noise, the control *belief* aided them in completing the task. In similar studies, perceived control benefited immune system functioning (Sieber et al., 1992). Outside of the laboratory, studies have shown that older residents in assisted living facilities, which are notorious for low control, lived *longer* and showed *better* health outcomes when given control over something as simple as watering a plant or choosing when student volunteers came to visit (Rodin & Langer, 1977; Schulz & Hanusa, 1978). In addition, feeling in control of a threatening situation can actually change stress hormone levels (Dickerson & Kemeny, 2004). Believing that you have control over your own behaviors can also have a positive influence on important outcomes like smoking cessation, contraception use, and weight management (Wallston & Wallston, 1978). When individuals do not believe they have control, they do not try to change. Self-efficacy is closely related to control, in that people with high levels of this trait believe they can complete tasks and reach their goals. Just as feeling in control can reduce stress and improve health, higher self-efficacy can reduce stress and negative health behaviors, and is associated with better health (O'Leary, 1985).

Social Relationships

Research has shown that the impact of social isolation on our risk for disease and death is similar in magnitude to the risk associated with smoking regularly (Holt-Lunstad, Smith, & Layton, 2010; House, Landis, & Umberson, 1988). In fact, the importance of social relationships for our health is so significant that some scientists believe our body has developed a physiological system that encourages us to seek out our relationships, especially in times of stress (Taylor et al., 2000). Social integration is the concept used to describe the number of social roles that you have (Cohen & Wills, 1985), as well as the lack of isolation. For example, you might be a daughter, a basketball team member, a Humane Society volunteer, a coworker, and a student. Maintaining these different roles can improve your health via encouragement from those around you to maintain a healthy lifestyle. Those in your social network might also provide you with social support (e.g., when you are under stress). This support might include emotional help (e.g., a hug when you need it), tangible help (e.g., lending you money), or advice. By helping to improve health behaviors and reduce stress, social relationships can have a powerful, protective impact on health, and in some cases, might even help people with serious illnesses stay alive longer (Spiegel, Kraemer, Bloom, & Gottheil, 1989).

Dispositions and Emotions: What's Risky and What's Protective?

Negative dispositions and personality traits have been strongly tied to an array of health risks. One of the earliest negative trait-to-health connections was discovered in the 1950s by two cardiologists. They made the interesting discovery that there were common behavioral and psychological patterns among their heart patients that were not present in other patient samples. This pattern included being competitive, impatient, hostile, and time urgent. They labeled it Type A Behavior. Importantly, it was found to be associated with *double* the risk of heart disease as compared with Type B Behavior (Friedman & Rosenman, 1959). Since the 1950s, researchers have discovered that it is the hostility and competitiveness components of Type A that are especially harmful to heart health (Iribarren et al., 2000; Matthews, Glass, Rosenman, & Bortner, 1977; Miller, Smith, Turner, Guijarro, & Hallet, 1996). Hostile individuals are quick to get upset, and this angry arousal can damage the arteries of the heart. In addition, given their negative personality style, hostile people often lack a heath-protective supportive social network.

Positive traits and states, on the other hand, are often health protective. For example, characteristics like positive emotions (e.g., feeling happy or excited) have been tied to a wide range of benefits such as increased longevity, a reduced likelihood of developing some illnesses, and better outcomes once you are diagnosed with certain diseases (e.g., heart disease, HIV) (Pressman & Cohen, 2005). Across the world, even in the most poor and

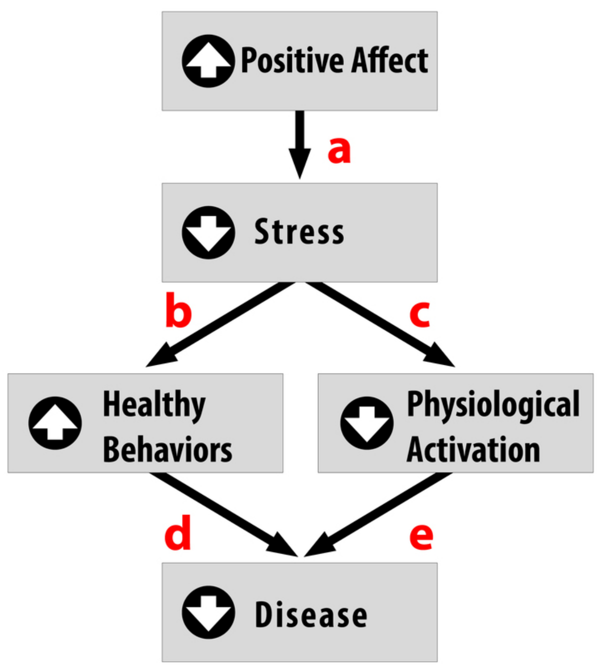


Figure 2. This figure illustrates one possible way that positive affect protects individuals against disease. Positive affect can reduce stress perceptions (a), thereby improving health behaviors (b) and lowering physiological stress responses (c) (e.g., decreased cardiovascular reactivity, lower stress hormones, non-suppressed immune activity). As a result, there is likely to be less incidence of disease (d, e). (Adapted from Pressman & Cohen, 2005)

underdeveloped nations, positive emotions are consistently tied to better health (Pressman, Gallagher, & Lopez, 2013). Positive emotions can also serve as the "antidote" to stress, protecting us against some of its damaging effects (Fredrickson, 2001; Pressman & Cohen, 2005; see Figure 2). Similarly, looking on the bright side can also improve health. Optimism has been shown to improve coping, reduce stress, and predict better disease outcomes like recovering from a heart attack more rapidly (Kubzansky, Sparrow, Vokonas, & Kawachi, 2001; Nes & Segerstrom, 2006; Scheier & Carver, 1985; Segerstrom, Taylor, Kemeny, & Fahey, 1998).

Stress Management

About 20 percent of Americans report having stress, with 18–33 year-olds reporting the highest levels (American Psychological Association, 2012). Given that the sources of our stress are often difficult to change (e.g., personal finances, current job), a number of interventions have been designed to help reduce the aversive responses to duress. For example, relaxation activities and forms of meditation are techniques that allow individuals to reduce their stress via breathing exercises, muscle relaxation, and mental imagery. Physiological arousal from stress can also be reduced via **biofeedback**, a technique where the individual is shown bodily information that is not normally available to them (e.g., heart rate), and then taught strategies to alter this signal. This type of intervention has even shown promise in reducing heart and hypertension risk, as well as other serious conditions (e.g., Moravec, 2008; Patel, Marmot, & Terry, 1981). But reducing stress does not have to be complicated! For example, exercise is a great stress reduction activity (Salmon, 2001) that has a myriad of health benefits.

The Importance Of Good Health Practices

As a student, you probably strive to maintain good grades, to have an active social life, and to stay healthy (e.g., by getting enough sleep), but there is a popular joke about what it's like to be in college: you can only pick two of these things (see Figure 3 for an example). The busy life of a college student doesn't always allow you to maintain all three areas of your life, especially during test-taking periods. In one study, researchers found that students taking exams were more stressed and, thus, smoked more, drank more caffeine, had less physical activity, and had worse sleep habits (Oaten & Chang, 2005), all of which could have detrimental effects on their health. Positive health practices are *especially* important in times of stress when your immune system is compromised due to high stress and the elevated frequency of exposure to the illnesses of your fellow students in lecture halls, cafeterias, and dorms.

Psychologists study both **health behaviors** and health habits. The former are behaviors that can improve or harm your health. Some examples include regular exercise, flossing, and

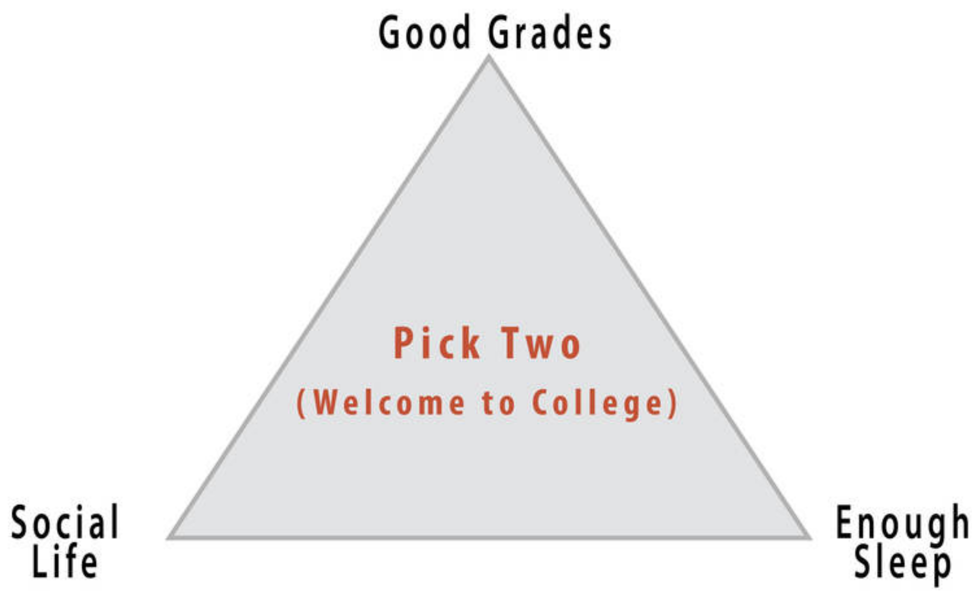


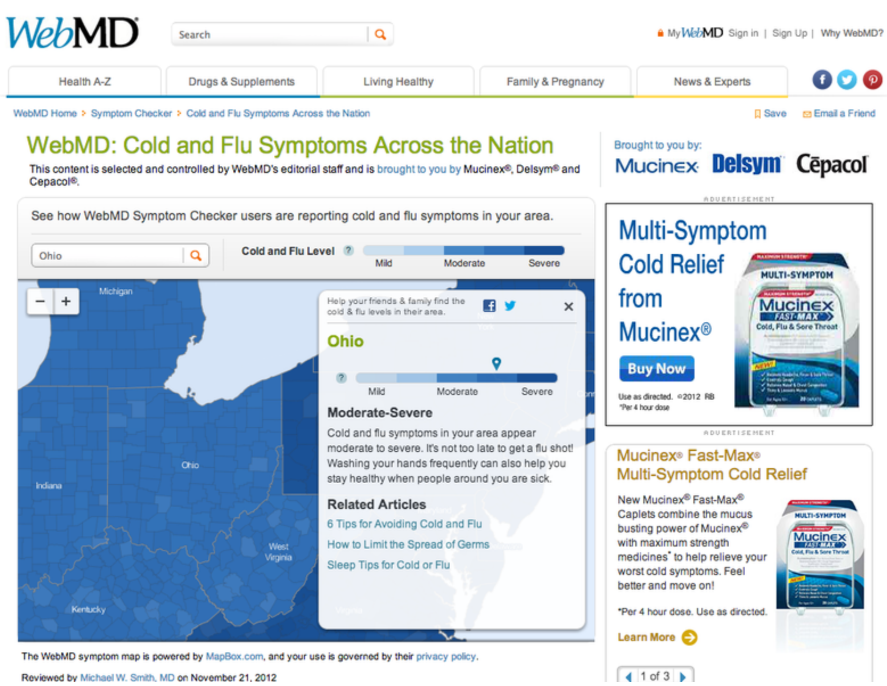
Figure 3: A popular joke about how difficult it is to stay balanced and healthy during college.

wearing sunscreen, versus negative behaviors like drunk driving, pulling all-nighters, or smoking. These behaviors become *habits* when they are firmly established and performed automatically. For example, do you have to think about putting your seatbelt on or do you do it automatically? Habits are often developed early in life thanks to parental encouragement or the influence of our peer group.

While these behaviors sound minor, studies have shown that those who engaged in more of these protective habits (e.g., getting 7–8 hours of sleep regularly, not smoking or drinking excessively, exercising) had fewer illnesses, felt better, and were less likely to die over a 9–12-year follow-up period (Belloc & Breslow 1972; Breslow & Enstrom 1980). For college students, health behaviors can even influence academic performance. For example, poor sleep quality and quantity are related to weaker learning capacity and academic performance (Curcio, Ferrara, & De Gennaro, 2006). Due to the effects that health behaviors can have, much effort is put forward by psychologists to understand *how* to change unhealthy behaviors, and to understand *why* individuals fail to act in healthy ways. Health promotion involves enabling individuals to improve health by focusing on behaviors that pose a risk for future illness, as well as spreading knowledge on existing risk factors. These might be genetic risks you are born with, or something you developed over time like obesity, which puts you at risk for Type 2 diabetes and heart disease, among other illnesses.

Psychology And Medicine

There are many psychological factors that influence medical treatment outcomes. For example, older individuals, (Meara, White, & Cutler, 2004), women (Briscoe, 1987), and those from higher socioeconomic backgrounds (Adamson, Ben-Shlomo, Chaturvedi, & Donovan, 2008) are all *more* likely to seek medical care. On the other hand, some individuals who need care might avoid it due to financial obstacles or preconceived notions about medical practitioners or the illness. Thanks to the growing amount of medical information online, many people now use the Internet for health information and 38% percent report that this influences their decision to see a doctor (Fox & Jones, 2009). Unfortunately, this is not always a good thing because individuals tend to do a poor job assessing the credibility of health information. For example, college-student participants reading online articles about HIV and syphilis rated a physician's article and a college student's article as *equally* credible if the participants said they were familiar with the health topic (Eastin, 2001). Credibility of health information often means how accurate or trustworthy the information is, and it can be influenced by irrelevant factors, such as the website's design, logos, or the organization's contact information (Freeman & Spyridakis, 2004). Similarly, many people post health questions on online, unmoderated forums where *anyone* can respond, which allows for the possibility of inaccurate information being provided for serious medical conditions by unqualified individuals.



While the Internet has increased the amount of medical information available to the public and created greater access, there are real concerns about how people are making decisions about their health based on that information. [Image: Mapbox, https://goo.gl/UNhmx5, CC BY 2.0, https://goo.gl/BRvSA7]

After individuals decide to seek care, there is also variability in the information they give their medical provider. Poor communication (e.g., due to embarrassment or feeling rushed) can influence the accuracy of the diagnosis and the effectiveness of the prescribed treatment. Similarly, there is variation following a visit to the doctor. While most individuals are tasked with a health recommendation (e.g., buying and using a medication appropriately, losing weight, going to another expert), not everyone *adheres* to medical recommendations (Dunbar-Jacob & Mortimer-Stephens, 2010). For example, many individuals take medications inappropriately (e.g., stopping early, not filling prescriptions) or fail to change their behaviors (e.g., quitting smoking). Unfortunately, getting patients to follow medical orders is not as easy as one would think. For example, in one study, over one third of diabetic patients failed to get proper medical care that would prevent or slow down diabetes-related blindness (Schoenfeld, Greene, Wu, & Leske, 2001)! Fortunately, as mobile technology improves, physicians now have the ability to monitor **adherence** and work to improve it (e.g., with pill bottles that monitor if they are opened at the right time). Even text messages are useful for improving treatment adherence and outcomes in depression, smoking cessation, and weight loss (Cole-Lewis, & Kershaw, 2010).

Being A Health Psychologist

Training as a clinical health psychologist provides a variety of possible career options. Clinical health psychologists often work on teams of physicians, social workers, allied health professionals, and religious leaders. These teams may be formed in locations like rehabilitation centers, hospitals, primary care offices, emergency care centers, or in chronic illness clinics. Work in each of these settings will pose unique challenges in patient care, but the primary responsibility will be the same. Clinical health psychologists will evaluate physical, personal, and environmental factors contributing to illness and preventing improved health. In doing so, they will then help create a treatment strategy that takes into account all dimensions of a person's life and health, which maximizes its potential for success. Those who specialize in health psychology can also conduct research to discover new health predictors and risk factors, or develop interventions to prevent and treat illness. Researchers studying health psychology work in numerous locations, such as universities, public health departments, hospitals, and private organizations. In the related field of **behavioral medicine**, careers focus on the application of this type of research. Occupations in this area might include jobs in occupational therapy, rehabilitation, or preventative medicine. Training as a health psychologist provides a wide skill set applicable in a number of different professional settings and career paths.

The Future Of Health Psychology

Much of the past medical research literature provides an incomplete picture of human health. "Health care" is often "illness care." That is, it focuses on the management of symptoms and illnesses as they arise. As a result, in many developed countries, we are faced with several health epidemics that are difficult and costly to treat. These include obesity, diabetes, and cardiovascular disease, to name a few. The National Institutes of Health have called for researchers to use the knowledge we have about risk factors to design effective interventions to reduce the prevalence of *preventable* illness. Additionally, there are a growing number of individuals across developed countries with *multiple* chronic illnesses and/or lasting disabilities, especially with older age. Addressing their needs and maintaining their quality of life will require skilled individuals who understand how to properly treat these populations. Health psychologists will be on the forefront of work in these areas.

With this focus on prevention, it is important that health psychologists move beyond studying risk (e.g., depression, stress, hostility, low socioeconomic status) in isolation, and move toward studying factors that confer resilience and protection from disease. There is, fortunately, a growing interest in studying the positive factors that protect our health (e.g., Diener & Chan, 2011; Pressman & Cohen, 2005; Richman, Kubzansky, Maselko, Kawachi, Choo, & Bauer, 2005) with evidence strongly indicating that people with higher positivity live longer, suffer fewer illnesses, and generally feel better. Seligman (2008) has even proposed a field of "Positive Health" to specifically study those who exhibit "above average" health—something we do not think about enough. By shifting some of the research focus to identifying and understanding these health-promoting factors, we may capitalize on this information to improve public health.

Innovative interventions to improve health are already in use and continue to be studied. With recent advances in technology, we are starting to see great strides made to improve health with the aid of computational tools. For example, there are hundreds of simple applications (apps) that use email and text messages to send reminders to take medication, as well as mobile apps that allow us to monitor our exercise levels and food intake (in the growing mobile-health, or m-health, field). These m-health applications can be used to raise health awareness, support treatment and compliance, and remotely collect data on a variety of outcomes. Also exciting are devices that allow us to monitor physiology in real time; for example, to better understand the stressful situations that raise blood pressure or heart rate. With advances like these, health psychologists will be able to serve the population better, learn more about health and health behavior, and develop excellent health-improving strategies that could be specifically targeted to certain populations or individuals. These leaps in equipment development, partnered with growing health psychology knowledge and exciting advances in neuroscience and genetic research, will lead health researchers and practitioners into an exciting new time where, hopefully, we will understand more and more about how to keep people healthy.

Outside Resources

App: 30 iPhone apps to monitor your health
http://www.hongkiat.com/blog/iphone-health-app/

Quiz: Hostility
http://www.mhhe.com/socscience/hhp/fahey7e/wellness_worksheets/wellness_worksheet_090.html

Self-assessment: Perceived Stress Scale
http://www.ncsu.edu/assessment/resources/perceived_stress_scale.pdf

Video: Try out a guided meditation exercise to reduce your stress
https://www.youtube.com/watch?v=dEzbdLn2bJc

Web: American Psychosomatic Society
http://www.psychosomatic.org/home/index.cfm

Web: APA Division 38, Health Psychology
http://www.health-psych.org

Web: Society of Behavioral Medicine
http://www.sbm.org

Discussion Questions

1. What psychological factors contribute to health?
2. Which psychosocial constructs and behaviors might help protect us from the damaging effects of stress?
3. What kinds of interventions might help to improve resilience? Who will these interventions help the most?
4. How should doctors use research in health psychology when meeting with patients?
5. Why do clinical health psychologists play a critical role in improving public health?

Vocabulary

Adherence
In health, it is the ability of a patient to maintain a health behavior prescribed by a physician. This might include taking medication as prescribed, exercising more, or eating less high-fat food.

Behavioral medicine
A field similar to health psychology that integrates psychological factors (e.g., emotion, behavior, cognition, and social factors) in the treatment of disease. This applied field includes clinical areas of study, such as occupational therapy, hypnosis, rehabilitation or medicine, and preventative medicine.

Biofeedback
The process by which physiological signals, not normally available to human perception, are transformed into easy-to-understand graphs or numbers. Individuals can then use this information to try to change bodily functioning (e.g., lower blood pressure, reduce muscle tension).

Biomedical Model of Health
A reductionist model that posits that ill health is a result of a deviation from normal function, which is explained by the presence of pathogens, injury, or genetic abnormality.

Biopsychosocial Model of Health
An approach to studying health and human function that posits the importance of biological, psychological, and social (or environmental) processes.

Chronic disease
A health condition that persists over time, typically for periods longer than three months (e.g., HIV, asthma, diabetes).

Control
Feeling like you have the power to change your environment or behavior if you need or want to.

Daily hassles
Irritations in daily life that are not necessarily traumatic, but that cause difficulties and repeated stress.

Emotion-focused coping
Coping strategy aimed at reducing the negative emotions associated with a stressful event.

General Adaptation Syndrome
A three-phase model of stress, which includes a mobilization of physiological resources phase, a coping phase, and an exhaustion phase (i.e., when an organism fails to cope with the stress adequately and depletes its resources).

Health
According to the World Health Organization, it is a complete state of physical, mental, and social well-being and not merely the absence of disease or infirmity.

Health behavior
Any behavior that is related to health—either good or bad.

Hostility
An experience or trait with cognitive, behavioral, and emotional components. It often includes cynical thoughts, feelings of emotion, and aggressive behavior.

Mind–body connection
The idea that our emotions and thoughts can affect how our body functions.

Problem-focused coping
A set of coping strategies aimed at improving or changing stressful situations.

Psychoneuroimmunology
A field of study examining the relationship among psychology, brain function, and immune function.

Psychosomatic medicine
An interdisciplinary field of study that focuses on how biological, psychological, and social processes contribute to physiological changes in the body and health over time.

Resilience
The ability to "bounce back" from negative situations (e.g., illness, stress) to normal functioning or to simply not show poor outcomes in the face of adversity. In some cases, resilience may lead to better functioning following the negative experience (e.g., post-traumatic growth).

Self-efficacy

The belief that one can perform adequately in a specific situation.

Social integration
The size of your social network, or number of social roles (e.g., son, sister, student, employee, team member).

Social support
The perception or actuality that we have a social network that can help us in times of need and provide us with a variety of useful resources (e.g., advice, love, money).

Stress
A pattern of physical and psychological responses in an organism after it perceives a threatening event that disturbs its homeostasis and taxes its abilities to cope with the event.

Stressor
An event or stimulus that induces feelings of stress.

Type A Behavior
Type A behavior is characterized by impatience, competitiveness, neuroticism, hostility, and anger.

Type B Behavior
Type B behavior reflects the absence of Type A characteristics and is represented by less competitive, aggressive, and hostile behavior patterns.

References

Adamson, J., Ben-Shlomo, Y., Chaturvedi, N., & Donovan, J. (2008). Ethnicity, socio-economic position and gender—do they affect reported health—care seeking behaviour? *Social Science & Medicine, 57*, 895–904.

American Psychological Association (2012). Stress in American 2012 [Press release]. Retrieved from http://www.apa.org/news/press/releases/stress/2012/generations.aspx

Belloc, N. B., & Breslow, L. (1972). Relationship of physical health status and health practices. *Preventive Medicine, 1*, 409–421.

Billings, A. G., & Moos, R. H. (1981). The role of coping responses and social resources in attenuating the stress of life events. *Journal of Behavioral Medicine, 4*, 139–157.

Breslow, L., & Enstrom, J. E. (1980). Persistence of health habits and their relationship to mortality. *Preventive Medicine, 9*, 469–483.

Briscoe, M. E. (1987). Why do people go to the doctor? Sex differences in the correlates of GP consultation. *Social Science & Medicine, 25*, 507–513.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology, 56*, 267–283.

Cohen, S., & Herbert, T. B. (1996). Health psychology: Psychological factors and physical disease from the perspective of human psychoneuroimmunology. *Annual Review of Psychology, 47*, 113–142.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98*, 310–357.

Cohen, S., Alper, C. M., Doyle, W. J., Treanor, J. J., & Turner, R. B. (2006). Positive emotional style predicts resistance to illness after experimental exposure to rhinovirus or influenza A virus. *Psychosomatic Medicine, 68*, 809–815.

Cohen, S., Janicki-Deverts, D., & Miller, G. E. (2007). Psychological stress and disease. *Journal of the American Medical Association, 298*, 1685–1687.

Cohen, S., Tyrrell, D. A., & Smith, A. P. (1991). Psychological stress and susceptibility to the common cold. *New England Journal of Medicine, 325*, 606–612.

Cole-Lewis, H., & Kershaw, T. (2010). Text messaging as a tool for behavior change in disease prevention and management. *Epidemiologic Reviews, 32*, 56–69.

Curcio, G., Ferrara, M., & De Gennaro, L. (2006). Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews, 10*, 323–337.

DeLongis, A., Folkman, S., & Lazarus, R. S. (1988). The impact of daily stress on health and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology, 54*, 486–495.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*, 355–391.

Dunbar-Jacob, J., & Mortimer-Stephens, M. (2001). Treatment adherence in chronic disease. *Journal of Clinical Epidemiology, 54*(12), S57–S60

Eastin, M. S. (2001). Credibility assessments of online health information: The effects of source expertise and knowledge of content. *Journal of Computer Mediated Communication, 6*.

Fox, S. & Jones, S. (2009). The social life of health information. *Pew Internet and American Life Project, California HealthCare Foundation*. Retrieved from http://www.pewinternet.org/Reports/2009/8-The-Social-Life-of-Health-Information.aspx

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218–226.

Freeman, K. S., & Spyridakis, J. H. (2004). An examination of factors that affect the credibility of online health information. *Technical Communication, 51*, 239–263.

Friedman, M., & Rosenman, R. (1959). Association of specific overt behaviour pattern with blood and cardiovascular findings. *Journal of the American Medical Association, 169*, 1286–1296.

Glass, D. C., & Singer, J. E. (1972). Behavioral aftereffects of unpredictable and uncontrollable aversive events: Although subjects were able to adapt to loud noise and other stressors in laboratory experiments, they clearly demonstrated adverse aftereffects. *American Scientist, 60*, 457–465.

Herman-Stabl, M. A., Stemmler, M., & Petersen, A. C. (1995). Approach and avoidant coping: Implications for adolescent mental health. *Journal of Youth and Adolescence, 24*, 649–665.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine, 7*(7), e1000316.

House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationships and health. *Science, 241*, 540–545.

Iribarren, C., Sidney, S., Bild, D. E., Liu, K., Markovitz, J. H., Roseman, J. M., & Matthews, K. (2000). Association of hostility with coronary artery calcification in young adults. *Journal of the American Medical Association, 283*, 2546–2551.

Kubzansky, L. D., Sparrow, D., Vokonas, P., & Kawachi, I. (2001). Is the glass half empty or half full? A prospective study of optimism and coronary heart disease in the normative aging study. *Psychosomatic Medicine, 63*, 910–916.

Matthews, K. A., Glass, D. C., Rosenman, R. H., & Bortner, R. W. (1977). Competitive drive, pattern A, and coronary heart disease: A further analysis of some data from the Western Collaborative Group Study. *Journal of Chronic Diseases, 30*, 489–498.

Meara, E., White, C., & Cutler, D. M. (2004). Trends in medical spending by age, 1963–2000. *Health Affairs, 23*, 176–183.

Miller, T. Q., Smith, T. W., Turner, C. W., Guijarro, M. L., & Hallet, A. J. (1996). Meta-analytic review of research on hostility and physical health. *Psychological Bulletin, 119*, 322–348.

Moravec, C. S. (2008). Biofeedback therapy in cardiovascular disease: rationale and research overview. *Cleveland Clinic Journal of Medicine, 75*, S35–S38.

Nes, L. S., & Segerstrom, S. C. (2006). Dispositional optimism and coping: A meta-analytic review. *Personality and Social Psychology Review, 10*, 235–251.

Oaten, M., & Cheng, K. (2005). Academic examination stress impairs self–control. *Journal of Social and Clinical Psychology, 24*, 254–279.

O'Leary, A. (1985). Self-efficacy and health. *Behaviour Research and Therapy, 23*, 437–451.

Patel, C., Marmot, M. G., & Terry, D. J. (1981). Controlled trial of biofeedback-aided behavioural methods in reducing mild hypertension. *British Medical Journal (Clinical research ed.), 282*, 2005–2008.

Pressman, S. D., & Cohen, S. (2005). Does positive affect influence health? *Psychological Bulletin, 131*, 925–971.

Pressman, S. D., Gallagher, M. W., & Lopez, S. J. (2013). Is the emotion-health connection a "first-world problem"? *Psychological Science, 24*, 544–549.

Richman, L. S., Kubzansky, L., Maselko, J., Kawachi, I., Choo, P., & Bauer, M. (2005). Positive emotion and health: Going beyond the negative. *Health Psychology, 24*, 422–429.

Rodin, J., & Langer, E. J. (1977). Long-term effects of a control-relevant intervention with the institutionalized aged. *Journal of Personality and Social Psychology, 35*, 897–902.

Rutter, M. (1985). Resilience in the face of adversity. *British Journal of Psychiatry, 147*, 598–611.

Salmon, P. (2001). Effects of physical exercise on anxiety, depression, and sensitivity to stress: A unifying theory. *Clinical Psychology Review, 21*(1), 33–61.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: assessment and implications of generalized outcome expectancies. *Health Psychology, 4*, 219–247.

Schoenfeld, E. R., Greene, J. M., Wu, S. Y., & Leske, M. C. (2001). Patterns of adherence to diabetes vision care guidelines: Baseline findings from the Diabetic Retinopathy Awareness Program. *Ophthalmology, 108*, 563–571.

Schulz, R., & Hanusa, B.H. (1978). Long-term effects of control and predictability-enhancing interventions: Findings and ethical issues. *Journal of Personality and Social Psychology, 36*, 1194–1202.

Segerstrom, S. C., Taylor, S. E., Kemeny, M. E., & Fahey, J. L. (1998). Optimism is associated with mood, coping, and immune change in response to stress. *Journal of Personality and Social Psychology, 74*, 1646–1655.

Seligman, M. E. P. (2008). Positive health. *Applied Psychology, 57*, 3–18.

Selye, H. (1946). The general adaptation syndrome and the diseases of adaptation. *Journal of Clinical Endocrinology, 6*, 117–230.

Sieber, W. J., Rodin, J., Larson, L., Ortega, S., Cummings, N., Levy, S., ... Herberman, R. (1992). Modulation of human natural killer cell activity by exposure to uncontrollable stress. *Brain, Behavior, and Immunity, 6*, 141–156.

Spiegel, D., Kraemer, H., Bloom, J., & Gottheil, E. (1989). Effect of psychosocial treatment on survival of patients with metastatic breast cancer. *The Lancet, 334*, 888–891.

Taylor, S. E. (2012) *Health psychology* (8th ed.). New York, NY: McGraw-Hill.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review, 107*, 411–429.

Twisk, J. W., Snel, J., Kemper, H. C., & van Mechelen, W. (1999). Changes in daily hassles and life events and the relationship with coronary heart disease risk factors: A 2-year longitudinal study in 27–29-year-old males and females. *Journal of Psychosomatic Research, 46*, 229–240.

Wallston, B. S., & Wallston, K. A. (1978). Locus of control and health: a review of the literature. *Health Education & Behavior, 6*, 107–117.

World Health Organization (2013). *Cardiovascular diseases*. Retrieved from http://www.who.int/mediacentre/factsheets/fs317/en/index.html

World Health Organization. (1946). *Preamble to the Constitution of the World Health Organization*. Retrieved from http://www.who.int/about/definition/en/print.html

Index

5α-reductase, 129
abducens nucleus, 271
ablate, 74
acetylcholine, 269
adaptations, 106
adherence, 679
adoption study, 185
Affective, 562
afferent nerve fibers, 266
A-fibers, 245
age identity, 478
aggression, 510
aggressive behavior, 131
agnosias, 91
agonists, 655
Agreeableness, 546
agreeableness, 530
allodynia, 250
altruism, 533
ambulatory assessment, 25
amygdala, 492, 567
analgesics, 249
anchor, 433
anhedonia, 620
animistic soul, 607
antagonists, 655
anterograde amnesia, 391
aphasias, 89
arcuate fasciculus, 91
aromatase, 130
arousal: cost–reward model, 532
aspartate, 266
asylums, 608
attributional styles, 626
autobiographical memory, 339
autobiographical narratives, 477
automatic process, 495
autonomic nervous systems, 88
availability heuristic, 512
average life expectancy, 480
basal ganglia, 71
behavioral genetics, 185
behavioral medicine, 679
Behaviorism, 55
biases, 430
binocular advantage, 203
biofeedback, 676
Biomedical Model of Health, 670
biopsychosocial model, 611
Biopsychosocial Model of Health, 670
blocking, 367
blood-oxygen-level-dependent (BOLD), 38
bounded awareness, 435
bounded ethicality, 435
bounded rationality, 430
brain stem, 69
Broca’s area, 89
bystander intervention, 526
callosotomy, 72
case studies, 74
categorize, 371
catharsis, 516
cathartic method, 610
central sulcus, 89
cerebellum, 69, 92
cerebral cortex, 70
cerebral hemispheres, 70
cerebrum, 69, 89

C-fibers, 246
chromosomal sex, 129
chronic disease, 669
chronic pain, 250
chronic stress, 626
chunk, 325
Chutes and Ladders, 453
cingulate gyrus, 91
classical, 361
classical conditioning, 323
cochlea, 217
cognitive psychology, 54
cohorts, 474
compensatory reflexes, 270
computerized axial tomography, 93
concrete operational reasoning stage, 449
concrete operations stage, 451
conditioned aversions and preferences, 233
conditioned compensatory responses, 366
conditioned response (CR), 362
conditioned stimulus (CS), 362
cones, 198
Conscientiousness, 546
consciousness, 52
conservation problems, 450
consolidated, 388
consolidation, 344
context, 369
continuous, 449
continuous distributions, 544
contralateral, 72
contrast, 199
contrast gain, 202
control, 673
converging evidence, 73
Convoy Model of Social Relations, 478
cortisol, 499
cost–benefit analysis, 528
C-pain or Aδ-fibers, 245
crossmodal phenomena, 293
crossmodal receptive fields, 291
crossmodal stimuli, 291
Cross-sectional studies, 475
Crowds, 464
Crystallized intelligence, 475
C-tactile, 246
cue overload principle, 347
cultural display rules, 583
cultural relativist, 606
cutaneous senses, 243
daily hassles, 672
dark adaptation, 202
day reconstruction method (DRM), 22
decay, 387
declarative memory, 391
deoxygenated hemoglobin, 38
depolarization, 41
depolarizes, 266
depth perception, 447
descending pain modulatory system, 248
Detection thresholds, 267
deviant peer contagion, 463
diary method, 21
dichotic listening, 402
differential susceptibility, 468
Diffuse optical imaging (DOI), 76
diffusion of responsibility, 528
directionally tuned, 268
discontinuous., 449
discriminative stimulus, 370
dissociative amnesia, 392
distinctiveness, 340
divided attention, 401
DNA methylation, 162
DNA methyltransferases (DNMTs), 162
double flash illusion, 296
drive state, 592

early adversity, 626
ecological momentary assessment, 20
ecological validity, 19
ectoderm, 87
EEG, 494
egoism, 533
electroencephalogram, 494
Electroencephalography (EEG), 76
electronically activated recorder, or EAR, 23
Emotion-focused coping, 673
empathic concern, 533
empathy–altruism model, 533
empirical methods, 8
empiricism, 50
encoding, 325, 339, 387
encoding specificity principle, 346
endocrine glands, 120
Endophenotypes, 647
endorphin, 250
engrams, 343
enzyme induction, 659
enzymes, 658
epigenetic, 146
Epigenetics, 161
epigenome, 161
episodic memory, 338
Error management theory (EMT), 112
ethics, 12
etiology, 606
eugenics, 54
event-related potentials, 93
event-related potentials (ERP), 644
evolution, 105
experience-sampling method, 20
external validity, 19
exteroception, 243
extinction, 369
Extraversion, 545, 546
facets, 548
factor analysis, 546
false memories, 313
fear conditioning, 366
feminization, 126
fight or flight response, 499
Five-Factor Model, 546
flashbulb memory, 55, 341
Fluid intelligence, 475
foils, 311
forebrain, 87
Foreclosure, 465
formal operational reasoning stage, 449
formal operations stage, 451
fornix, 91
framing, 434
frontal cortex, 565
frontal lobe, 71, 89
full-cycle psychology, 27
functional magnetic resonance imaging, 496
Functional magnetic resonance imaging (fMRI), 75, 644
functional neuroanatomy, 498
functionalism, 53
g, 416
gamma-aminobutyric acid, 269
gaze stability, 270
gene, 160
General Adaptation Syndrome, 672
generalize, 18
genotype, 162
Gestalt psychology, 54
global subjective well-being, 479
globus pallidus, 91
glutamate, 266
goal-directed, 373
gonadal sex, 129
grandiosity, 620
gray matter, 72, 92

gustation, 230
gyri, 70
gyrus, 70, 89
habit, 373
habituation, 323
hair cells, 265
health, 670
health behaviors, 674, 676
Hedonic well-being, 480
helpfulness, 530
helping, 526
hemoglobin, 38
heritability coefficient, 187
heterogeneity, 474
heuristics, 431
HEXACO model, 550
hippocampus, 91
histone acetyltransferases (HATs) and histone deacetylases (HDACs), 163
histone modifications, 163
homeostasis, 592
Homo habilis, 86
Homo sapiens, 86
homophily, 463
Hormones, 120
hormones, 499
hostile attribution bias, 514
hostile expectation bias, 514
hostile perception bias, 514
hostility, 675
Humorism, 607
hyperpolarization, 41
hyperpolarizes, 266
hypersomnia, 620
hypothalamic-pituitary-adrenal (HPA) axis, 499
hypothalamus, 91, 565, 596
hypotheses, 8
hysteria, 607
identical twins, 162
Identity achievement, 465
Identity diffusion, 465
immunocytochemistry, 92
implicit learning, 323
implicit memory, 323
incidental learning, 325
individual differences, 54
Information processing theories, 447
ingroup, 500
inhibitory functioning, 476
instrumental, 363
integrated, 288
intelligence, 415
intentional learning, 325
interaural level differences (ILDs), 221
interaural time differences (ITD), 220
interference, 389
internal validity, 19
interoception, 243
interpersonal, 577
intersexual selection, 108
intra- and inter-individual differences, 474
intrapersonal, 577
intrasexual competition, 107
introspection, 52
invasive, 36
IQ, 418
kin selection, 531
lateral geniculate nucleus, LGN, 91
lateral inhibition, 201
lateral rectus muscle, 271
lateral sulcus, 89
lateral vestibulo-spinal tract, 272
lateralized, 72
law of effect, 364
lesion studies, 93
lesions, 36, 74, 492
lexical hypothesis, 546

Life course theories, 474
Life span theories, 474
limbic system, 71, 91
limited capacity, 400
linguistic analyses, 26
longitudinal studies, 475
magnetic resonance imaging, 93
magnification factor, 89
maladaptive, 606
masculinization, 126
maternal behavior, 122
McGurk effect, 295
mechanically gated ion channels, 266
medial prefrontal cortex, 496
medial temporal lobes, 391
medial vestibulo-spinal tract, 272
medulla oblongata, 92
memory traces, 343
mentalizing, 497
mesmerism, 610
Metabolism, 658
metabolites, 68
Metacognition, 327
mind–body connection, 671
misinformation effect, 309, 345
mnemonic devices, 350
mock witnesses, 311
Moratorium, 465
motor cortex, 71
multimodal, 288
multimodal perception, 288
multimodal phenomena, 293
multisensory convergence zones, 290
multisensory enhancement, 289
myelin, 72
natural selection, 106
Nature, 447
negative state relief model, 532
neural crest, 87
neural impulse, 51
neural induction, 87
neural plasticity, 41
neuroblasts, 87
neuroendocrinology, 499
neuroepithelium), 87
neuroscience, 563
neuroscience methods, 36
Neuroticism, 546
neurotransmitter, 655
neurotransmitters, 120, 266
nociception, 243
nociceptors, 243
nomenclature, 69
nonassociative learning, 323
noninvasive, 37
normed, 419
noxious stimuli, 250
nucleus accumbens, 566
numerical magnitudes, 453
Nurture, 447
object permanence task, 449
observational learning, 375
occipital lobe, 70, 91
oculomotor nuclei, 270
oculomotor nucleus, 272
olfaction, 230
omnivores, 230
Openness, 546
operant, 363
operant conditioning, 323, 363
Opponent Process theory of color, 204
orthonasal olfaction, 233
other-oriented empathy, 530
otoconia, 268
otolith receptors, 264
outgroup, 500
overconfident, 433
oxygenated hemoglobin, 38

oxytocin, 120, 145
pain, 242
parasympathetic nervous system (PNS), 42
parasympathetic nervous systems., 88
Parental behavior, 133
parietal lobe, 71, 90
paternal, 134
Pavlovian conditioning, 361
perceptual learning, 323
periaqueductal gray, 567
peripheral nervous system, 41
personal distress, 533
personality, 544
Personality traits, 544
person-situation debate, 552
phantom pain, 252
Pharmacokinetics, 656
phenotype, 161
phonemic awareness, 452
photo spreads, 311
photoactivation, 198
phrenology, 73
Piaget's stage theory, 446
pinna, 217
placebo effect, 249
pluralistic ignorance, 527
polypharmacy, 662
pons, 92
positron, 40
Positron Emission Tomography, 93
Positron emission tomography (PET), 75
prediction error, 368
preoperational reasoning stage, 449
preoptic area, 568, 597
preparedness, 368
primary auditory cortex, 292
primary motor cortex, 89
primary somatosensory cortex, 90
primary visual cortex, 202, 292
Principle of Inverse Effectiveness, 289
Problem-focused coping, 673
processing speed, 475
progesterone, 134
progestin, 120
prolactin, 120
proprioceptive, 270
prosocial behavior, 530
prosocial personality orientation, 530
psychoactive drugs, 655
psychogenic, 606
psychological adaptations, 109
Psychological control, 463
psychometric approach, 475
psychomotor agitation, 620
psychoneuroimmunology, 671
psychophysics, 51
Psychophysiological methods, 36
psychosomatic medicine, 671
psychotropic drug, 656
punishers, 364
Punishment, 518
qualitative changes, 448
quantitative changes, 448
quantitative genetics, 186
quantitative law of effect, 371
recall, 475
receptive field, 291
receptors, 122
Reciprocal altruism, 531
recoding, 341
recognition, 476
reinforcer, 363
reinforcer devaluation effect, 372
relational aggression, 513
renewal effect, 369
resilience, 672
retardation, 620
retrieval, 339, 388

retroactive interference, 345
Retrograde amnesia, 391
retronasal olfaction, 233
reward value, 597
rods, 198
rostrocaudal, 87
rubber hand illusion, 295
satiation, 596
schemata, 313
scholar-practitioner model, 57
scientist-practitioner model, 57
selective attention, 401
Self-efficacy, 674
self-interest is bounded, 435
self-perceptions of aging, 478
Semantic memory, 338
semicircular canals, 264
sensitization, 323
Sensitization, 251
sensorimotor stage, 449
sensory modalities, 287
set point, 593
sex determination, 129
sexual differentiation, 128
sexual selection, 107
shadowing, 402
simulation, 495
social and cultural, 577
social brain, 643
Social categorization, 493
Social integration, 674
Social Learning Theory, 375
social models, 375
social network, 478
social referencing, 581
social support, 501, 674
social touch hypothesis, 247
social zeitgeber, 628
Sociocultural theories, 447
socioeconomic status, 623
Socioemotional Selectivity Theory, 478
somatic nervous system, 88
somatogenic, 606
somatosensory (body sensations) cortex, 71
somatosensory cortex, 245
somatotopically organized, 245
spatial principle of multisensory integration, 291
spatial resolution, 38, 74
spina bifida, 87
split-brain patients, 72
spontaneous recovery, 369
standardized, 418
stereocilia, 265
stereotype threat, 423
stereotypes, 493
stimulus control, 370
storage, 339
stress, 672
Stress, 499
stressors, 672
stria terminalis, 568
structuralism, 52
subcortical, 70
Subjective age, 477
successful aging, 481
suicidal ideation, 620
sulci, 70
sulcus, 70, 89
superadditive effect of multisensory integration, 289
superior temporal sulcus, 497
supernatural, 606
sympathetic nervous system, 88, 499
sympathetic nervous system (SNS), 42
synapse, 655
syndrome, 612

synesthesia, 206
System 1, 436
System 2, 436
systematic observation, 8
target cells, 122
taste aversion conditioning, 366
temporal lobe, 71, 89
temporal parietal junction, 497
temporal resolution, 38, 74
temporally graded retrograde amnesia, 391
testosterone, 120
thalamus, 91, 567
the perception of collisions between two circles (called “balls”), 297
theories, 8
tip-of-the-tongue phenomenon, 55
torsion, 272
traitement moral, 609
Transcranial direct current stimulation (tDCS), 74
Transcranial magnetic stimulation (TMS), 74
transduction, 91, 243
transfer-appropriate processing, 327
Trephination, 606
Trichromacy theory, 204
twin studies, 186
tympanic membrane, 217
Type A Behavior, 675
Type B Behavior, 675
unconditioned response (UR), 362
unconditioned stimulus (US), 362
unimodal, 287
unimodal components, 289
unimodal cortex, 292
vagus nerve, 145
vasopressin, 147
vergence angle, 272
vestibular compensation, 269
vestibular efferents, 269
vestibular system, 263
vestibulocochlear nerve, 266
vestibuloocular reflex, 270
vestibulo-ocular reflex, 206
vicarious reinforcement, 377
violence, 510
visual cortex, 567
visual hemifield, 72
voltage, 39
Wernicke’s area, 91
What pathway, 203
Where-and-How pathway, 203
white coat hypertension, 24
white matter, 72, 92
willpower is bounded, 435
working memory, 89, 324, 475