



Fooled by the brain: Re-examining the influence of neuroimages



N.J. Schweitzer^{a,*}, D.A. Baker^a, Evan F. Risko^b

^a Arizona State University, United States

^b University of Waterloo, Canada

ARTICLE INFO

Article history:

Received 5 November 2012

Revised 29 July 2013

Accepted 6 August 2013

Available online 14 September 2013

Keywords:

Neuroimages

Bias

Judgment

Persuasion

ABSTRACT

A series of highly-cited experiments published in 2008 demonstrated a biasing effect of neuroimages on lay perceptions of scientific research. More recent work, however, has questioned this bias, particularly within legal contexts in which neuroscientific evidence is proffered by one of the parties. The present research moves away from the legal framework and describes five experiments that re-examine this effect. Experiments 1 through 4 present conceptual and direct replications of some of the original 2008 experiments, and find no evidence of a neuroimage bias. A fifth experiment is reported that confirms that, when laypeople are allowed multiple points of reference (e.g., when directly comparing neuroimagery to other graphical depictions of neurological data), a neuroimage bias can be observed. Together these results suggest that, under the right conditions, a neuroimage might be able to bias judgments of scientific information, but the scope of this effect may be limited to certain contexts.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The judgments individuals' make can be influenced by a number of factors ostensibly unrelated to that judgment (Tversky & Kahneman, 1974). Understanding these factors provide clues regarding how individuals make judgments and decisions and can have wide-reaching societal implications, for example, by allowing policy makers to institute policies that minimize the negative impact of these biases (e.g., in legal, medical, and economic contexts). A salient example of such research, published in *Cognition*, emerged recently with the demonstration that the presence of neuroimages altered individuals' judgments of scientific credibility (McCabe & Castel, 2008). This finding has had a striking impact in cognitive science and beyond as evidenced, for example, by research in the legal domain where "neuroevidence" has the potential to bias legal deci-

sion makers (Appelbaum, 2009; Compton, 2010; Roskies, Schweitzer, and Saks, 2013; Vincent, 2011). This subsequent research, however, has largely failed to corroborate the original findings. In the present investigation, we return to the original context (rather than a legal context) in order to explore this discrepancy further.

1.1. The impact of neuroimages

McCabe and Castel (2008) presented three experiments demonstrating that neuroscientific information has the potential to unduly influence decision-making. Their participants were given brief articles that summarized a fictitious neuroscience study. The articles were manipulated to contain a neuroimage, a non-neuroimage graphical representation of the experimental data (e.g., a bar graph or a topographical map of the brain), or no image at all. In each experiment the authors reported findings that suggested the articles accompanied by the brain image were judged more favorably than the control articles. The authors concluded, "...there is, indeed, something special about the

* Corresponding author. Address: Arizona State University, MC3051, Phoenix, AZ 85069-7100, United States. Tel.: +1 602 543 8133.

E-mail address: njs@asu.edu (N.J. Schweitzer).

brain images with respect to influencing judgments of scientific credibility” (p. 350). This result was corroborated by Weisberg, Keil, Goodstein, Rawson, and Gray (2008) who presented neuroscience students and laypeople with descriptions of psychological phenomena. When these descriptions included neuroscientific language both groups were more satisfied with the description of the phenomena, even though the neuroscientific language was designed to be irrelevant. Further corroboration comes from a recent study by Keehner, Mayberry, and Fischer (2011) who presented their participants with neurological information in several graphical formats (e.g., a topographical map of the brain, a 3-D color representation of the brain) and found that certain highly-realistic depictions of the brain made the accompanying passages more convincing relative to the less-realistic images. Most recently, Ikeda, Kitagami, Takahashi, Hattori, and Ito (in press) demonstrated both an effect of neuroimages on credibility judgments and on metacomprehension judgments. That is, individuals felt like they understood a text passage better when an image of the brain was included compared to no image or a bar graph.

If neuroimages hold the sort of persuasive power suggested by McCabe and Castel (2008), Keehner et al. (2011) and Ikeda et al. (in press), then this would raise a number of potential concerns about the use of such information in decision-making contexts. This concern is particularly salient in legal contexts where presenting such neuroimages may very well be considered prejudicial if admitted into a trial. Indeed, in response to such findings, legal scholars’ existing concerns over neuroimaging (Baskin, Edersheim, & Price, 2007; Dumit, 1999; Pratt, 2005; Rose, 2000) began to intensify (Appelbaum, 2009; Brown & Murphy, 2010; Compton, 2010; Erickson, 2010; Vincent, 2011). These concerns appeared to be well supported. For example, Gurley and Marcus (2008) examined the impact of neuroimaging on mock jurors’ willingness to find a criminal defendant Not Guilty by Reason of Insanity (NGRI). When their participants were given a written piece of expert testimony that included a structural neuroimage, a greater number of participants rendered NGRI verdicts—based on their belief in the findings of the neurological evidence—compared to the same expert testimony that lacked neuroimage evidence.

1.2. Reexamining the impact of neuroimages

Provided both the theoretical and practical significance of these findings, a number of subsequent experiments assessed the impact of neuroimages on judgment in legal contexts and elsewhere. Interestingly, almost no evidence emerged that supported the notion that brain images have any impact on judgment in these contexts. For example, Schweitzer et al. (2011) used a large nationally-representative sample to assess the impact of neuroimage evidence when offered in support of a *mens rea* defense in a criminal trial. Across four different experiments varying the severity of the charged crime, the type of neuroimage used, and the visual impact (“glitziness”) of the image, evidence that contained neuroimages was in no way more persuasive than similar evidence that did not contain neuroimages (see also

Greene & Cahill, 2011; Schweitzer & Saks, 2011 for similar results in a legal context). It is important to note that, while these experiments focused on legal decisions, mediating variables that measured the credibility of the specific scientific evidence associated with the neuroimage (more akin to measures used in the original studies) were also collected, and were also unaffected by the presence of neuroimagery. In addition, these failures to replicate do not seem limited to the legal context. For example, Gruber and Dickerson (2012) paired popular media science articles with neuroimages, sci-fi images, and fantasy-like artistic renderings. In line with the legal work on neuroimages, no persuasive impact of neuroimages were found. Similarly, Michael, Newman, Vuorre, Cumming, and Garry (2013) sought to replicate the McCabe and Castel (2008) findings. Across multiple attempts with various samples and presentation modalities, Michael et al. were unable to replicate the findings reported in McCabe and Castel’s third experiment, finding that perceptions of a newspaper-type article describing a scientific finding were not influenced by the inclusion of an accompanying neuroimage. Finally, Hook and Farah (2013), examined neuroimagery’s influence on judgments of research summaries, again finding no persuasive impact, and no relation between neuroimage effects and a person’s belief in mind–brain dualism.

2. The present experiments

The conflicting findings within this body of literature are puzzling. Looking specifically at the methodology of the earlier studies that purportedly support a neuroimage bias explains only part of the discrepancy. For example, Weisberg et al. (2008) focused on neuroscience language generally without specifically testing the impact of neuroimagery. Gurley and Marcus (2008) confounded the presentation of neuroimages with additional expert evidence, making it impossible to parse the contribution of the neuroimage. Lastly, Keehner et al. (2011) presented neuroimages in a repeated-measures design, opening the door to contrast effects in which neuroimages may be persuasive only when other points of reference are given. Making a similar claim with respect to McCabe and Castel (2008), however, is more difficult and thus the potential reason for the incongruity is not as clear. It could be the specific task used in the studies, that the dependent construct was not properly operationalized, or that the effects found were simply idiosyncratic or spurious. In five experiments, we endeavored to understand this remaining discrepancy by attempting to re-examine the neuroimage bias by conceptually and directly replicating McCabe and Castel’s (2008) findings and then attempting to identify the conditions most likely to elicit a neuroimage effect.

2.1. Experiment 1

In Experiment 1 we conducted a conceptual replication of McCabe and Castel (2008) – hereafter noted as “M&C” – using a web-based experiment that presented participants with one of three different scenarios describing a purported link between a particular mental condition and a

resulting behavior. The three scenarios were chosen to create a spectrum of potential situations in which a neuroimage might be used as supporting evidence. The first described an idiopathic brain defect that would cause a woman to behave erratically; the second suggested that exposure to a neurotoxin would cause a brain abnormality that causes an individual to behave erratically; and the third indicated that brain damage caused by past drug use would lead to reduced functionality in a person's brain; all scenarios are included as [Appendix A](#). In support of these assertions, each scenario was accompanied by either a graph or a neuroimage, illustrated in [Appendix B](#).

Participants ($N = 175$; obtained online via Amazon Mechanical Turk) were randomly assigned to one of the six conditions in this 3×2 design and were then presented with a series of questions designed to index the extent to which the participants were convinced by the evidence (listed in [Table 1](#)).¹ These items were adapted from [Schweitzer et al. \(2011\)](#) in which they were found to predict judgments of guilt in a criminal case. These items—for example, “To what extent do you believe the neurotoxin can cause a brain defect?”—were measured on a seven-point scale, from “not at all” to “completely”.

The four DV items were averaged into a single composite measure of believability (Cronbach's $\alpha = .63$), and a 2 (Image) \times 3 (Scenario) between-subjects analysis of variance was conducted on this DV. While there was a main effect of the Scenario, $F(2, 169) = 24.66$, $p < .001$, $\eta_p^2 = .23$, the interaction was not significant, $F(2, 169) = 1.15$, $p = .61$, $\eta_p^2 = .006$, and, critically, there was no effect of the image, $F(1, 169) = 1.15$, $p = .28$, $\eta_p^2 = .007$. These results are illustrated in [Fig. 1](#).²

To further explore the data, individual univariate ANOVAs compared individual DV items between the graph and neuroimage version of each of the three scenarios. None of the individual DV items significantly differed between the two image conditions (see [Table 1](#)). In sum, the presence of a neuroimage did not make any scenario more or less believable than when a graph containing similar information was presented.

2.2. Experiment 2

After finding no evidence of a neuroimage effect in our first experiment, we undertook a second conceptual replication that more closely resembled the original M&C procedures. We created two brief (fictitious) news articles that described the findings of a recent piece of brain research: the first described a purported relation between retinal damage and brain function, and the second described a link

between brain abnormalities in childhood and subsequent criminal behavior (similar to M&C's original materials; summaries listed in [Appendix A](#)). As with the first experiment, the articles were accompanied with either a graph or a neuroimage as supporting evidence; however, in this case we used the original color images (bar graph and brain image) that appeared in [Fig. 1a](#) of M&C.

A sample of participants ($N = 222$; obtained online via Amazon Mechanical Turk) were randomly assigned to one of the conditions and were asked a series of six questions designed to measure the persuasiveness of the research. The dependent items were adapted from earlier work ([Schweitzer & Saks, 2012](#)) in which the items were found to predict judgments of guilt within a court case that revolved around the veracity of a piece of scientific evidence. As with Experiment 1, the items (listed in [Table 1](#)) asked participants the extent to which they agreed with various statements, with responses falling on a seven-point scale ranging from “not at all” to “completely”. A 2 (Article Topic) \times 2 (Image) ANOVA was conducted on the composite of our six DVs (Cronbach's $\alpha = .87$), finding that while our “retinal damage” article was seen as more persuasive than the “criminal” article, $F(1, 218) = 12.09$, $p < .001$, $\eta_p^2 = .053$, there was no main effect of the image manipulation $F(1, 218) = .12$, $p = .72$, $\eta_p^2 = .001$, nor any interaction $F(1, 218) = 1.47$, $p = .23$, $\eta_p^2 = .007$. These results are illustrated in [Fig. 2](#).³ In addition, the image manipulation did not influence any of the individual DVs (p s range from .28 to .93; individual tests listed in [Table 1](#)). Once again, we failed to find any difference in the reported persuasive impact of scientific research as a function of the presence of neuroimaging.

2.3. Experiment 3

In Experiment 3 we attempted a close conceptual replication of M&C's third experiment. First, we obtained the original BBC News article used in M&C Experiment 3 ([BBC News, 2005](#)). This article, entitled “Can Brain Scans Detect Criminals”, summarizes an experiment in which fMRI techniques were used to predict when participants were lying vs. telling the truth. Further, we again used the graph and brain image presented in [Fig. 1a](#) of M&C's original manuscript. A sample of 300 participants, obtained via Amazon Mechanical Turk,⁴ (56.6% male; mean age

¹ In this and subsequent experiments, we utilized multiple dependent measure items to assess the extent to which participants judged our scenarios. To simplify and facilitate visualizing these data we created a single composite measure made from averaging these items. Following the composite tests, and to ensure a thorough analysis, each item is also tested separately.

² As a guard against inattention, participants who took <45 s (as based on pilot testing of a reasonably minimal length for completion) were not included in the analysis, although this cutoff did not substantively affect the results. For example, using the full set of participants, the significance level of Image on the Composite DV changes from $p = .28$ to $p = .39$. The mean time for completion was 83 s ($SD = 47$ s).

³ As in Experiment 1, participants who completed their ratings in fewer than 45 s were excluded, although, once again, this cutoff did not substantively affect the significance of any of the tests. For example, using the full set of responses, the significance level of the Image manipulation on the Composite DV changes from $p = .72$ to $p = .64$. Mean completion time = 120 s ($SD = 92$ s).

⁴ To be particularly thorough in guarding against Type II errors, we excluded 32 participants who failed the manipulation check (answering fewer than 75% of the items correctly) and/or completed the experiment in fewer than 200 s (as based on pilot testing of a reasonably minimal length for completion), leaving a final sample of 268. Mean completion time was 351 s ($SD = 173$ s). As with previous experiments, we re-analyzed these data using different cutoffs for attention, finding no substantive differences. For example, using the full set of participants changes the Conclusion DV (which was the closest to significance) from $p = .16$ to $p = .20$. Using a more stringent cutoff of only participants who scored 100% on the recall quiz resulted in a significance of $p = .19$.

Table 1

List of means (standard deviations) and effect sizes for DVs across all experiments.

	Brain	Graph	Control	Cohen's <i>d</i> [95% CI]
Experiment 1 (overall)	<i>N</i> = 81	<i>N</i> = 94		
To what extent do you believe the [subject] has a brain defect?	5.75 (1.39)	5.89 (1.23)		-.108 [-.300, .085]
...the [subject's] problems are caused by a brain defect?	5.27 (1.61)	5.32 (1.46)		-.033 [-.258, .193]
...the [subject] might be a danger to herself or others?	5.59 (1.18)	5.62 (1.15)		-.026 [-.197, .146]
...the [subject] will be in control of her actions?	3.22 (1.25)	3.46 (1.28)		-.191 [-.377, -.004]
Composite DV	5.32 (1.05)	5.34 (0.87)		-.021 [-.162, .120]
Experiment 1 (idiopathic)	<i>N</i> = 22	<i>N</i> = 39		
To what extent do you believe the [subject] has a brain defect?	6.55 (0.80)	6.44 (0.72)		.149 [-.036, .334]
...the [subject's] problems are caused by a brain defect?	5.95 (1.25)	6.05 (1.00)		-.093 [-.363, .178]
...the [subject] might be a danger to herself or others?	6.05 (1.00)	5.95 (1.07)		.097 [-.161, .355]
...the [subject] will be in control of her actions?	2.68 (0.99)	2.90 (1.19)		-.199 [-.476, .078]
Composite DV	5.97 (0.72)	5.88 (0.69)		.131 [-.042, .304]
Experiment 1 (neurotoxin)	<i>N</i> = 27	<i>N</i> = 32		
To what extent do you believe the [subject] has a brain defect?	5.33 (1.49)	5.00 (1.39)		.234 [-.127, .594]
...the [subject's] problems are caused by a brain defect?	6.04 (0.94)	5.63 (1.07)		.412 [.158, .666]
...the [subject] might be a danger to herself or others?	5.52 (1.12)	5.41 (1.13)		.099 [-.183, .382]
...the [subject] will be in control of her actions?	3.30 (1.30)	3.72 (1.37)		-.319 [-.655, .016]
Composite DV	5.40 (0.79)	5.08 (0.83)		.401 [.197, .615]
Experiment 1 (drug use)	<i>N</i> = 32	<i>N</i> = 23		
To what extent do you believe the [subject] has a brain defect?	5.56 (1.44)	6.21 (1.00)		-.519 [-.850, -.188]
...the [subject's] problems are caused by a brain defect?	4.16 (1.67)	3.65 (1.30)		.340 [-.056, .736]
...the [subject] might be a danger to herself or others?	5.34 (1.29)	5.35 (1.19)		-.008 [-.332, .316]
...the [subject] will be in control of her actions?	3.53 (1.29)	4.04 (0.93)		-.450 [-.750, -.151]
Composite DV	4.81 (1.19)	4.79 (0.69)		.020 [-.243, .283]
Experiment 2 (overall)	<i>N</i> = 109	<i>N</i> = 113		
To what extent do you feel you understand this study?	4.95 (1.28)	5.04 (1.43)		-.067 [-.244, .111]
...feel the researcher did a "good" job?	4.98 (1.19)	4.99 (1.38)		-.008 [-.177, .161]
...feel the scientific reasoning in the article made sense?	5.08 (1.34)	4.91 (1.50)		.120 [-.066, .306]
...believe the study's findings?	4.63 (1.36)	4.58 (1.48)		.035 [-.151, .222]
...believe this is a "scientific" study?	5.17 (1.48)	5.14 (1.56)		.020 [-.179, .219]
...feel this article was well written?	5.35 (1.21)	5.16 (1.33)		.150 [-.017, .317]
Composite DV	5.03 (1.00)	4.97 (1.17)		.055 [-.087, .198]
Experiment 2 (retina)	<i>N</i> = 55	<i>N</i> = 55		
To what extent do you feel you understand this study?	5.00 (1.39)	5.09 (1.38)		-.066 [-.322, .191]
...feel the researcher did a "good" job?	5.27 (1.01)	5.09 (1.32)		.155 [-.063, .372]
...feel the scientific reasoning in the article made sense?	5.44 (1.12)	5.22 (1.45)		.171 [-.069, .411]
...believe the study's findings?	5.18 (1.16)	4.89 (1.33)		.235 [.003, .466]
...believe this is a "scientific" study?	5.67 (1.39)	5.33 (1.35)		.250 [-.003, .504]
...feel this article was well written?	5.60 (1.03)	5.20 (1.19)		.363 [.157, .569]
Composite DV	5.36 (0.83)	5.14 (1.05)		.235 [.059, .410]
Experiment 2 (criminal)	<i>N</i> = 54	<i>N</i> = 58		
To what extent do you feel you understand this study?	4.91 (1.19)	4.98 (1.49)		-.052 [-.301, .196]
...feel the researcher did a "good" job?	4.69 (1.29)	4.90 (1.45)		-.154 [-.406, .098]
...feel the scientific reasoning in the article made sense?	4.72 (1.46)	4.62 (1.51)		.068 [-.205, .341]
...believe the study's findings?	4.07 (1.33)	4.29 (1.57)		-.152 [-.420, .116]
...believe this is a "scientific" study?	4.67 (1.40)	4.97 (1.74)		-.191 [-.482, .100]
...feel this article was well written?	5.09 (1.34)	5.12 (1.45)		-.022 [-.278, .235]
Composite DV	4.69 (1.05)	4.81 (1.27)		-.104 [-.318, .111]
Experiment 3	<i>N</i> = 133	<i>N</i> = 134		
To what degree do you agree or disagree that the title "Brain Scans Can Detect Criminals" is a good summary of the results?	3.27 (1.55)	3.13 (1.61)		.089 [-.100, .278]
...with the conclusion that brain imaging can be used as a lie detector?	4.02 (1.52)	3.78 (1.47)		.131 [-.018, .340]
To what extent do you feel the researcher did a "good" job?	4.89 (1.23)	4.78 (1.44)		.082 [-.078, .242]

...feel that the scientific reasoning in the article made sense?	4.99 (1.23)	4.90 (1.37)		.069 [–.086, .225]
...believe the study's findings?	4.38 (1.40)	4.22 (1.47)		.112 [–.060, .283]
...believe this is a "scientific" study?	5.00 (1.53)	4.97 (1.52)		.020 [–.162, .202]
...feel this article was well written?	5.18 (1.34)	5.01 (1.37)		.126 [–.036, .288]
Composite DV (items 3–7)	4.99 (0.98)	4.87 (1.07)		.117 [–.005, .240]
Experiment 4	N = 74	N = 67	N = 68	Brain vs. Graph Brain vs. Control
To what degree do you agree or disagree that the title "Brain Scans Can Detect Criminals" is a good summary of the results?	2.24 (0.79)	2.12 (0.75)	2.19 (0.81)	.157 [.030, .283]
To what extent to you agree or disagree with the conclusion that brain imaging can be used as a lie detector?	2.69 (0.68)	2.46 (0.77)	2.59 (0.72)	.063 [–.069, .194] .321 [.202, .439]
To what extent do you feel the researcher did a "good" job?	3.74 (0.80)	3.78 (0.87)	3.76 (0.72)	.144 [.030, .258] –.049 [–.185, .087] –.027 [–.150, .097]
...feel that the scientific reasoning in the article made sense?	3.81 (0.77)	3.66 (1.05)	3.90 (0.74)	.165 [.015, .315]
...believe the researcher's findings?	3.57 (0.78)	3.31 (1.00)	3.44 (0.92)	–.120 [–.243, .003] .339 [.165, .514] .147 [.001, .293]
...feel that the research described in the article was scientific?	3.73 (0.87)	3.67 (0.96)	3.71 (0.90)	.066 [–.083, .216]
I understood this article	4.24 (0.68)	4.37 (0.65)	4.41 (0.60)	.023 [–.121, .167] –.213 [–.322, .105] –.283 [–.388, –.178]
Composite DV (Items 3–7)	3.82 (0.56)	3.75 (0.67)	3.84 (0.54)	.115 [.014, .215] –.037 [.126, .053]
Experiment 5 (Block 1)	N = 81		N = 78	
I understood this summary	4.68 (0.47)		4.28 (1.06)	.494 [.368, .620]
I feel the researchers described in the summary did a good job	3.38 (1.20)		3.59 (1.18)	–.178 [–.361, .006]
I feel the scientific reasoning in the summary made sense	3.59 (1.23)		3.67 (1.28)	–.064 [–.258, .130]
I believe the researchers' findings	3.62 (1.12)		3.73 (0.96)	–.106 [–.267, .055]
I feel that the research described in the summary was "scientific"	3.16 (1.25)		3.22 (1.17)	–.050 [–.237, .137]
I agree with the researchers' conclusions	3.58 (1.12)		3.67 (1.04)	–.084 [–.251, .083]
Composite DV	3.67 (0.88)		3.69 (0.92)	–.022 [–.161, .117]
Experiment 5 (Block 2)	N = 78		N = 81	
I understood this summary	4.47 (0.75)		4.49 (0.64)	–.029 [–.136, .079]
I feel the researchers described in the summary did a good job	3.65 (1.07)		3.11 (1.31)	.453 [.268, .639]
I feel the scientific reasoning in the summary made sense	3.69 (1.09)		3.10 (1.21)	.515 [.337, .693]
I believe the researchers' findings	3.73 (0.96)		3.30 (1.21)	.395 [.226, .564]
I feel that the research described in the summary was "scientific"	3.41 (1.05)		2.83 (1.31)	.491 [.307, .674]
I agree with the researchers' conclusions	3.73 (0.94)		3.20 (1.24)	.484 [.313, .654]
Composite DV	3.78 (0.82)		3.34 (0.94)	.501 [.365, .638]

33.0 years; representing 45 US states; 45.7% of the sample holding a bachelor's degree or higher) were randomly assigned to view the BBC News article accompanied by either a neuroimage or a graph. Following the article, the participants answered the same two DV items posed by M&C, along with the same six additional measures of persuasiveness used in Experiment 2. In addition to the items used in Experiments 1 and 2, we further included several manipulation check items (a quiz of the materials), demographic items, and a subset of items from the Need for Cognition scale (Cacioppo, Petty, & Kao, 1984). This would, in part, allow us to identify if the neuroimage effects found by earlier researchers were present in a subset of the sample (as would be indicated by interactions between our IV and the demographic/NFC measures).

Looking first at the two DVs used in the original M&C experiment, we found no significant difference between the two conditions on the extent to which participants felt the title of the article was a good summary of the results $F(1, 266) = .56, p = .45, \eta_p^2 = .002$, nor on the extent to which the participants agreed with the conclusion of the article, $F(1, 266) = 2.02, p = .16, \eta_p^2 = .008$. Next, as with Experiment 2, an ANOVA comparing the conditions the composite score created from our six persuasiveness items (but not the M&C items; Cronbach's $\alpha = .85$) produced a nonsignificant result, $F(1, 265) = .832, p = .36, \eta_p^2 = .003$. These results are illustrated in Fig. 3. Further, individual univariate tests indicated no significant differences on any of the individual DVs (all $ps > .30$; results listed in Table 1). In addition, none of the individual difference

variables (age, gender, education, Need for Cognition) interacted with the image manipulation.

2.4. Experiment 4

In our third experiment we used the original McCabe and Castel materials comparing neuroimages to a graphical representation of the same information. However, in the M&C Experiment 3, neuroimages were compared to a no-image control condition, and so in this fourth experiment we added a condition containing no graphical information, resulting in an exact replication of M&C. We further made our replication authentic by including a “Criticism” IV as in M&C, in which we created two versions of the news article: one containing commentary from an expert who was critical of the research findings (which is present in the actual BBC article) and another version that excluded this criticism.

Participants ($N = 210$, 59% male, mean age = 28.5) obtained through Amazon Mechanical Turk were randomly assigned to conditions within a 3 (Image: Brain, Graph, or No Image) $\times 2$ (Criticism: Present or Absent) between-subjects design. Participants read the BBC news article which was once again followed by the same two DV items posed by M&C, along with a five-item quiz about the article they had read.⁵

A 3×2 between-subjects ANOVA was conducted on the first M&C item—the extent to which the participants agreed that the title was a good summary of the results. Consistent with M&C, we found no significant effect of the image manipulation $F(2, 204) = .640$, $p = .53$, $\eta_p^2 = .006$, nor an interaction, $F(2, 204) = .784$, $p = .42$, $\eta_p^2 < .008$; however, we further found no main effect of criticism, $F(1, 204) = 1.77$, $p = .18$, $\eta_p^2 = .009$. Critically, using uncorrected (LSD) paired comparison tests (to minimize the chance of a type II error) we found no significant difference between the Brain and Text Only Image conditions ($p = .64$). Turning to the second M&C item, another 3×2 between-subjects ANOVA indicated a marginal effect of image, $F(2, 204) = 2.39$, $p = .09$, $\eta_p^2 = .023$. Uncorrected paired comparison tests, however, showed that the brain did not significantly differ from the No Image condition used by M&C ($p = .38$); rather, the Brain condition yielded marginally more positive judgments than the Graph ($p = .06$). Consistent with M&C, we found a main effect of the criticism manipulation in which the article version containing criticism ($M = 2.47$) yielded significantly lower ratings than did the version containing no criticism ($M = 2.70$), $F(1, 204) = 6.16$, $p = .01$, $\eta_p^2 = .029$. No interaction was present, $F(2, 204) = .47$, $p = .63$, $\eta_p^2 = .005$.

Looking at the composite measure of our persuasiveness items (Cronbach's $\alpha = .76$), no main effect of image,

$F(2, 203) = .35$, $p = .70$, $\eta_p^2 = .003$, criticism, $F(1, 203) = .41$, $p = .52$, $\eta_p^2 = .002$, or interaction were found, $F(2, 203) = .58$, $p = .56$, $\eta_p^2 = .006$ nor did any of the individual DVs produce a significant difference (results listed in Table 1 and illustrated in Fig. 4).⁶

2.5. Experiment 5

Across our first four experiments, we did not identify any evidence of a neuroimage bias. Then why do these neuroimage effects intermittently appear? One possibility may lie in the format of the neuroimages used within the experiments: Keehner et al. (2011) describe the “ideal” neuroimage as a highly-realistic, 3-D, color image. While the experiments that demonstrate non-effects used various formats, none quite reached the maximum level of “glitziness” described by Keehner et al. Another possibility lies in the design of the experiments themselves: Several previous commentators (Schweitzer et al., 2011; Michael et al., *in press*) have speculated that the repeated-measures design used in experiments such as Weisberg et al. (2008) and Keehner et al. (2011) may facilitate neuroscience/neuroimage effects by giving participants a basis for comparison (which, by virtue of counterbalancing, are averaged into a single effect). The experiments finding non-effects have been between-subjects designs. Neuroimage effects may also depend on the quality of the information that the images accompany. For example, Weisberg et al. (2008) found that neuroscience information was particularly persuasive when it accompanied relatively weak or faulty arguments. Finally, it may be the case that the mere mention or framing of a scientific finding in neuroscientific terms may produce somewhat of a ceiling effect, preventing any additional influence by an accompanying neuroimage.

While our previous experiments have not produced evidence of bias, it remains an open question whether or not certain contexts exist in which a neuroimage could indeed bias judgment. The various contextual variables outlined above reflect potential avenues to pursue in moving toward a more nuanced understanding of the potential influence of a neuroimage. To begin this effort, in our final experiment, we combined a repeated measures design, passages containing weak arguments and no neuroscience language, and a realistic 3-D color neuroimage (compared to a text only control). Together these conditions produce seemingly the ideal environment to find a neuroimage effect. The lack of a neuroimage effect under these conditions would make it difficult to sustain the notion that neuroimages can bias judgment in any meaningful fashion. However, finding a neuroimage effect under these conditions would at least suggest that contexts exist where such an effect would emerge thus laying open the effort to identify those contexts and the mechanisms supporting neuroimage effects in those contexts.

⁵ 86.7% of participants scored at least four out of five correct answers on the quiz. Interestingly, recall scores differed as a function of the Image manipulation, with recall scores on the Graph condition being significantly lower than those on the Brain and Text Only conditions. As such, we did not exclude any participants on the basis of recall scores so as not to bias the results of our primary analysis. Mean completion time = 293 s (SD = 130 s). In addition, gender, age, and education were initially entered as factors or covariates in our models and did not substantively influence our findings, so these demographic measures are ignored for this experiment.

⁶ As with previous experiments, no substantive differences were produced when using different cutoffs for participant attention. For example, using our same cutoff from Experiment 3 (75%) changed the significance of the Image manipulation on the Conclusion DV (the strongest effect) from $p = .09$ to $p = .22$.

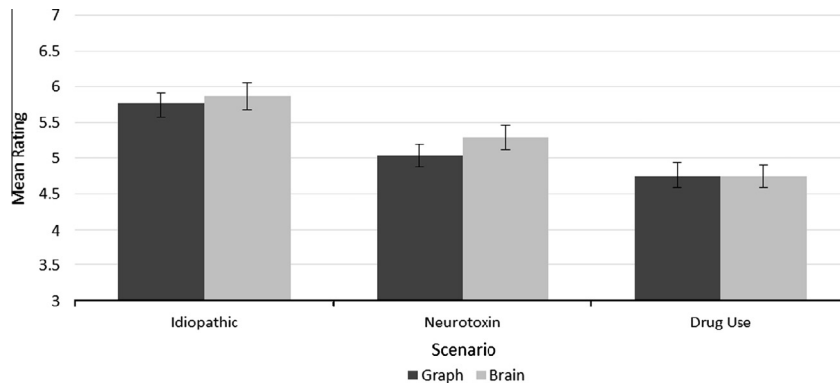


Fig. 1. Mean believability (± 1 S.E.) ratings by scenario and image, Experiment 1.

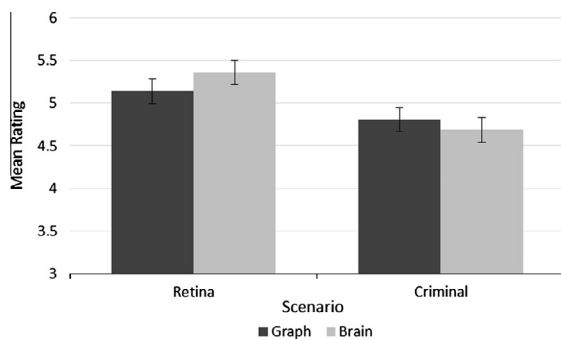


Fig. 2. Mean composite believability score (± 1 S.E.) by scenario and image, Experiment 2.

Participants obtained from Amazon Mechanical Turk ($N = 161$, 63% male, mean age = 30) completed this experiment. We utilized two brief summaries describing a piece of scientific research: The first summary was taken directly from Weisberg et al. (2008) (hereafter, the “knowledge” summary) and the second summary described a flawed experiment showing a link between a new drug and headaches used in Schweitzer and Saks (2012; hereafter the “headache” summary). Both summaries used faulty reasoning and had no neuroscience. Both summaries are included as Appendix A. Paired with each summary was either a color representation of a 3-dimensional fMRI image (similar to Keehner et al.’s (2011) most persuasive image; see Appendix B) or no image. Participants viewed and judged both summaries. The image-summary pairing was randomized, as was the order of presentation. Following each scenario, participants were given a set of six items measuring the extent to which the participant believed or was persuaded by the research described in the summary, rated on a 1–7 scale.

If the repeated measure format is indeed a driver of the effects found in Weisberg et al. (2008) and Keehner et al. (2011) we expect that any neuroimage effects would be found only in the second block (i.e., the second of the two “trials” and after being exposed to a passage) and not the first (i.e., the first of the two “trials” and before being exposed to any other passage). Looking only at the

data from the first block, we conducted a 2 (Image: Brain or No Image) \times 2 (Summary: Headache or Knowledge) between-subjects ANOVA on our composite measure of perceived credibility (Cronbach’s $\alpha = .89$). Despite the use of weak research scenarios, high-impact images, and no neuroscience language, the image manipulation did not yield a significant effect, $F(1, 155) = .19$, $p = .66$, $\eta_p^2 = .001$. A main effect of summary was found in that the Knowledge summary ($M = 3.92$) was rated as significantly more credible than the Headache summary ($M = 3.40$), $F(1, 155) = 14.78$, $p < .001$, $\eta_p^2 = .087$. In addition, an interaction was present, $F(1, 155) = 4.84$, $p = .03$, $\eta_p^2 = .030$. Looking within this interaction we find no significant effect of image within the Knowledge summary $F(1, 81) = .62$, $p = .43$, $\eta_p^2 = .007$, but a marginal effect within the Headache summary in which the brain image produces lower ratings ($M = 3.22$) than the No Image condition ($M = 3.58$), $F(1, 74) = 2.74$, $p = .10$, $\eta_p^2 = .036$. Tests on the individual DVs (listed in Table 1) produced a nearly identical pattern.

Turning to the second block (Cronbach’s $\alpha = .90$), we again found that the Knowledge summary ($M = 3.75$) was perceived to be significantly more credible than the Headache summary ($M = 3.38$), $F(1, 155) = 6.46$, $p = .01$, $\eta_p^2 = .040$. Critically, however, we found an effect of the Image manipulation with the scenario containing the brain image ($M = 3.78$), rated as significantly more credible than the summary containing No Image ($M = 3.34$), $F(1, 155) = 9.41$, $p = .003$, $\eta_p^2 = .056$. No interaction was present, $F(1, 155) = .480$, $p = .49$, $\eta_p^2 = .003$. Once again, tests on the individual DVs (listed in Table 1) produced a nearly identical pattern. These results are illustrated in Fig. 5 below.⁷

3. Meta-analysis

Finally, to maximize statistical power in an effort to uncover any neuroimage effects that might be present in our data, we combined the data from Experiments 1, 2, 3, and 4 to allow for a Brain vs. Graph comparison on our Compos-

⁷ Once again, using various timing cutoffs did not substantively change the results. For example, a cutoff of 150 s excluded the quickest 20% of the participants and changed the significance of the image effect in Block 1 from $p = .66$ to $p = .58$, and in Block 2 the significance remained constant at $p = .003$. Mean completion time = 242 s ($SD = 133$ s).

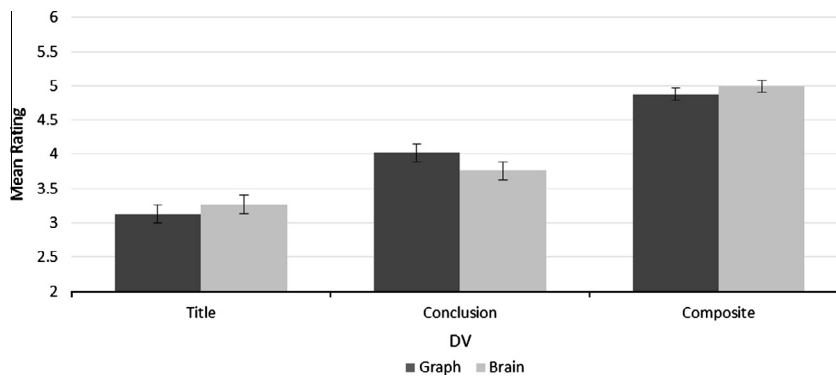


Fig. 3. Mean ratings by DV and image (± 1 SE). “Title” and “Conclusion” are original items used in McCabe and Castel (2008). “Composite” is the composite measure of our six items.

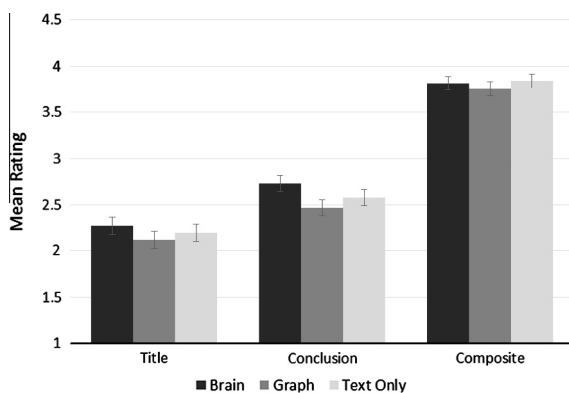


Fig. 4. Mean ratings by DV and image (± 1 SE), Experiment 4. “Title” and “Conclusion” are original items used in McCabe and Castel (2008). “Composite” is the composite measure of our six items.

ite DV of persuasiveness items (which was present in each of those experiments). Across 882 participants, a 2 (Brain vs. Graph) \times 4 (Experiment) ANOVA revealed no significant difference between the Brain ($M = 4.85$) and Graph ($M = 4.82$) conditions, $F(1, 874) = .797$, $p = .37$, $\eta_p^2 = .001$. Additionally, we combined the data from Experiment 4 and Block 1 of Experiment 5 ($N = 301$) and conducted a 2 (Experiment) \times 2 (Brain vs. Text Only) ANOVA on our composite item. Consistent with our earlier findings, no effect of image was found ($M_{\text{brain}} = 3.74$; $M_{\text{control}} = 3.76$) $F(1, 297) = .087$, $p = .77$, $\eta_p^2 = .001$.

4. Discussion

Across five experiments with over 1000 demographically-diverse participants, we were generally unable to find any influence of neuroimages on judgment across a range of scenarios. This included scenarios and dependent variables that were conceptually similar and identical to those used by McCabe and Castel (2008). These results corroborate the numerous demonstrations of a lack of a biasing influence of neuroimages in legal and other domains (Greene & Cahill, 2011; Gruber & Dickerson, 2012; Michael et al., 2013; Schweitzer & Saks, 2011; Schweitzer et al., 2011).

More importantly, the experiments reported here suggest that neuroimage effects are likely not present in a variety of different situations. For example, our first experiment revolved around medical and epidemiological decision-making and used three different scenarios that proposed a link between physiological, environmental, and toxicological causes and a specific neuro-behavioral defect. Across our second, third, and fourth experiments, we used a task that mirrors how the lay public often encounters science in the form of three different general-media “news articles” that described some new scientific finding. Our final experiment used two research abstracts as they might appear if read directly from a journal article.⁸ While some of these tasks are similar or identical to the M&C task, combined with the previous findings using legal decision-making tasks, it is becoming clear that neuroimages do not hold the sort of broad persuasive power as once suggested.

Although we were unable to replicate McCabe and Castel’s (2008) findings either conceptually or directly, this does not necessarily indicate that the original study was flawed; however, it does suggest that the biasing impact of neuroimagery is far less than was originally thought and, in fact, may be negligible. One potential explanation for these failures to replicate is that, as people have become more exposed to neuroscience over the past 5 years, the persuasive punch of neuroimagery has dulled, and what might have been legitimate effects in 2008 simply no longer exist. Unfortunately, testing this hypothesis would be quite difficult. In particular, it would be necessary to separate the effect of the novelty of neuroimagery from its substantive meaning—while pictures of brain function may no longer have the “wow” factor that they once had, as we learn more about the biological basis of human behavior, that single picture may take on additional meaning. In addition, both the Keehner et al. (2011) and Ikeda et al. (in press) have found effects of neuroimages and are, of course, more recent. Another possible factor in our lack inability to find a neuroimage effect involves our choice of sample and the modality of our experimental paradigm. While much of the work on the neuroscience

⁸ Though we are certain that academic researchers would be highly unlikely to evaluate a journal article by simply reading the abstract and glancing at the images.

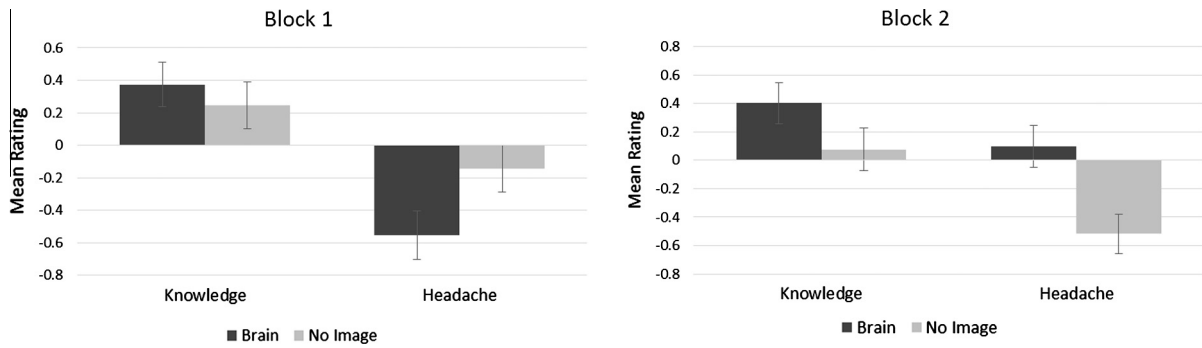


Fig. 5. Mean ratings (± 1 SE) of the Composite DV across scenario, image, and block.

bias utilized college-student samples and paper-and-pencil materials, the work described in this paper uses online community samples with the materials presented via computer. While these methodological differences certainly exist, it is unclear how they would account for the differences in results. In addition, other recent research on the neuroimage bias (e.g., Baker, Schweitzer, Risko, & Ware, *in press*; Michael et al., 2013) has used various participant groups (including college samples) and presentation methods (including computer presentation combined with paper and pencil response), while still finding no evidence of bias.

While the emerging skepticism about the persuasiveness of neuroimages is well justified, the present results also demonstrate that a neuroimage effect (akin to the original) does exist in at least some contexts. Namely, we were able to find a neuroimage effect when we used stimuli that combined the non-neuroscience language, weak arguments, and high-impact images described in Weisberg et al. (2008) and Keehner et al. (2011). However, even in this putatively optimal set of circumstances, the neuroimage effect occurred only after participants had multiple points of reference in a repeated measures design. Thus, the critical contextual variable appears to be the opportunity to (explicitly or implicitly) compare passages with and without an image of the brain in order for the brain image to exert any influence. In other words, a neuroimage bias might depend on the task involving a relative versus absolute judgment. Importantly, this discovery is not without practical consequences. For example, in legal contexts a jury is quite likely to be exposed to multiple sources of information and the situation would thus be more akin to a relative judgment task than to an absolute one.

Together the present results clearly define the limiting conditions on the influence neuroimagery on judgment across a variety of contexts. Critically, the present results also point in a new direction, ripe for future research, aimed at understanding what contexts within the identified limits might be sensitive to the biasing impact of neuroimages.

Appendix A

A.1. Research summaries used in Experiment 1

A.1.1. Idiopathic

Doctors have been examining the case of a woman who suddenly developed severe behavioral problems. Though

she had lived an otherwise normal and healthy life, shortly after her 34th birthday, the woman's behavior became erratic and aggressive. The woman would easily become upset and enraged, and would often yell at family members, friends, and strangers. The woman's problems grew increasingly worse until 1 day she was arrested after assaulting a stranger who had happened to bump into her on the sidewalk. After being examined by a team of doctors and psychologists, including undergoing an MRI scan of the woman's brain, the doctors concluded that she had a major defect in the left frontal lobe of her brain. A large portion of the brain tissue had essentially died, and was no longer functioning.

A.1.2. Neurotoxin

Scientists have recently been examining the effects of certain neurotoxins (poisons that target the brain and ner-

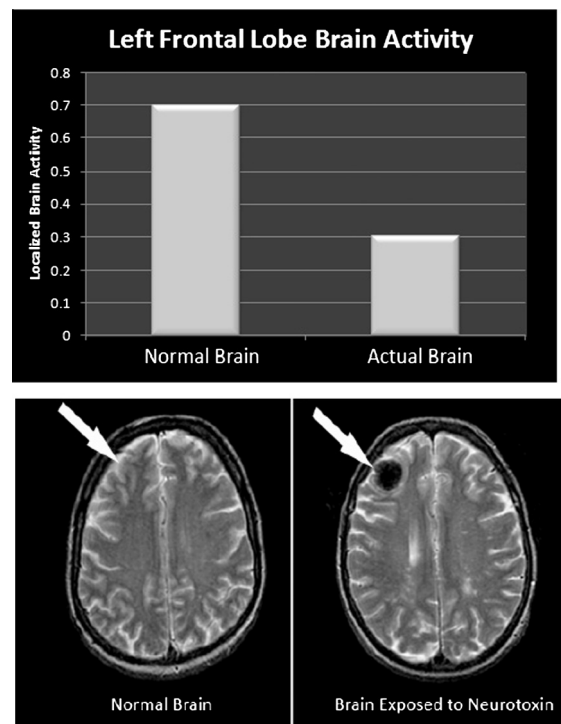


Fig. B1. Images used in Experiment 1 (presented to participants in black and white).

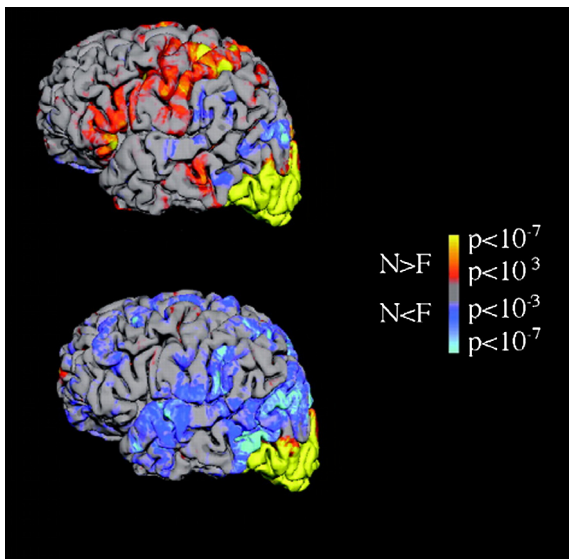


Fig. B2. Color 3-D neuroimage used in Experiment 5.

vous system) on human behavior. One such neurotoxin is known to cause sudden behavioral changes, making people become erratic and highly aggressive. One potential reason for these behaviors would be that the neurotoxin damages the left frontal lobe of the brain, which is responsible for planning and self-control. One individual who was recently exposed to this neurotoxin was given an MRI scan of the brain, and it was found that the left frontal lobe was damaged and had highly reduced function.

A.1.3. Past drugs

A recent study of the long-term effects of certain types of drug abuse involved conducting brain scans of drug abusers to assess the function of the left frontal lobe of their brains. The study found that many of the very worst of the long-term abusers suffered from damage and reduced functionality of their brains.

A.2. Research summaries used in Experiment 2

A.2.1. Brain abnormalities in childhood and subsequent criminal behavior

Many researchers believe that violent tendencies have a biological basis and that tests and brain imaging can pick them up. Prof Adam Blaine, a neurologist with substantial experience in diagnosing brain injuries, argues that abnormal physical brain make-up could be a cause of criminality. His studies have shown that psychopaths and criminals have significant brain abnormalities in the prefrontal cortex, which regulates and controls emotion and social behavior. He also investigated whether these abnormalities could be detected and treated in early childhood. Blaine conducted a study using magnetic resonance imaging (MRI) scan to examine the brains of more than 100 convicted criminals who had psychopathic tendencies or other anti-social behaviors that lead to criminal activities. He also used data from more than 9000 twins from

the Twins Early Development Study, a survey of twins born in England and Wales between 1994 and 1996. Assessments of callous unemotional traits and conduct problems were based on teacher questionnaires when the children were seven, nine and 12. Information was taken from parents when the children were as young as four. Blaine found there was a correlation between risk factors at a young age (seven) and bad behavior at an older age (twelve). The brain scans of these children also indicated a change in brain structure for those at higher risk for bad behavior that was consistent with the criminals. He acknowledged the ethical implications of treating children before they had done anything wrong, but argued that “biological” causes of crime could not be ignored.

A.2.2. Retinal damage and brain function

Looking at the back of the eye may offer insight into the health of someone’s brain, according to the US researchers. A study, published in the journal *Neurology*, linked damage to the retina with declining brain function. Researchers believe issues with the blood supply may be damaging both the eye and the brain. The eye condition the researchers were looking at was retinopathy, which is common in patients with Type 2 diabetes or high blood pressure. Damage to the retina can eventually lead to blindness. Scientists followed 511 women, who were 65 or older, for a decade. Nine were diagnosed with retinopathy. Those with the eye condition tended to have low scores in tests of brain function, including memory and abstract reasoning exams. Dr. Simon Riley, Alzheimer’s Research US, used magnetic resonance imaging (MRI) scans that revealed more areas of damaged brain tissue, ischemic lesions, in those with retinopathy. While there was no suggestion of dementia in the patients, brain decline can be an early sign of the disease. Riley points out that this could be very useful if simple eye screening could give an early indication that people might be at risk of problems with their brain health and functioning. The study adds to evidence linking vascular health to cognitive decline.

A.3. Research abstracts used in Experiment 5

A.3.1. “Knowledge” summary (taken directly from Weisberg et al. (2008) – No neuroscience language, weak conclusion)

Researchers created a list of facts that about 50% of people knew. Subjects in this experiment read the list of facts and had to say which ones they knew. They then had to judge what percentage of other people would know those facts. Researchers found that the subjects responded differently about other people’s knowledge of a fact when the subjects themselves knew that fact. If the subjects did know a fact, they said that an inaccurately large percentage of others would know it, too. For example, if a subject already knew that Hartford was the capital of Connecticut, that subject might say that 80% of people would know this, even though the correct answer is 50%. The researchers call this finding “the curse of knowledge”. The researchers claim that this “curse” happens because subjects make more mistakes when they have to judge the knowledge of others. People are much better at judging what they themselves know.

A.3.2. “Headache” summary (Pilot materials used in Schweitzer & Saks, 2012)

A study was conducted in order to determine whether a new drug causes people to get headaches. In order to test this, researchers handed out samples of the drug to 100 people. The researchers asked the people to record how often they took the drug, and also how many headaches they had. After 30 days, the researchers discovered that the more often people took the drug, the more headaches they had. The researchers concluded that the new drug causes people to get headaches.

Appendix B

See Figs. B1 and B2.

References

- Appelbaum, P. S. (2009). Through the glass darkly: Functional neuroimaging evidence enters the courtroom. *Psychiatric Services*, 60, 21–23. <http://dx.doi.org/10.1176/appi.ps.60.1.21>.
- Baker, D. A., Schweitzer, N. J., Risko, E. F., & Ware, J. (in press). Visual attention and the neuroimage bias. *PLOS One*.
- Baskin, J. H., Edersheim, J. G., & Price, B. H. (2007). Is a picture worth a thousand words? Neuroimaging in the courtroom. *American Journal of Law & Medicine*, 33(2–3), 239–269.
- BBC News (2005). *Can Brain Scans Detect Criminals?* <http://news.bbc.co.uk/2/hi/uk_news/4268260.stm>.
- Brown, T., & Murphy, E. (2010). Through a scanner darkly: Functional neuroimaging as evidence of criminal defendant's past mental states. *Stanford Law Review*, 62, 1119–1208.
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306–307.
- Compton, S. E. (2010). Note: Not guilty by reason of neuroimaging: The need for cautionary jury instructions for neuroscience evidence in criminal trials. *Vanderbilt Journal of Entertainment and Technology Law*, 12, 333–354.
- Dumit, J. (1999). Objective Brains, Prejudicial Images. *Science in Context*, 12(1), 173–201.
- Erickson, S. K. (2010). Blaming the brain. *Minnesota Journal of Law, Science & Technology*, 11, 27–77.
- Greene, E., & Cahill, B. S. (2011). Effects of neuroimaging evidence on mock juror decision making. *Behavioral Sciences & the Law*, 30(3), 280–296. <http://dx.doi.org/10.1002/bsl>.
- Gruber, D., & Dickerson, J. A. (2012). Persuasive images in popular science: Testing judgments of scientific reasoning and credibility. *Public Understanding of Science*. <http://dx.doi.org/10.1177/0963662512454072>.
- Gurley, J. R., & Marcus, D. K. (2008). The effects of neuroimaging and brain injury on insanity defenses. *Behavioral Sciences & the Law*, 26(1), 85–97. <http://dx.doi.org/10.1002/bsl.797>.
- Hook, C. J., & Farah, M. J. (2013). Look Again: Effects of Brain Images and Mind–Brain Dualism on Lay Evaluations of Research. *Journal of Cognitive Neuroscience*, 25(9), 1397–1405. <http://dx.doi.org/10.1162/jocn.a.00407>.
- Ikedo, K., Kitagami, S., Takahashi, T., Hattori, Y., & Ito, Y. (in press). Neuroscientific information bias in metacomprehension: The effect of brain images on metacomprehension judgment of neuroscience research. *Psychonomic Bulletin & Review*. <http://dx.doi.org/10.3758/s13423-013-0457-5>.
- Keehner, M., Mayberry, L., & Fischer, M. H. (2011). Different clues from different views: the role of image format in public perceptions of neuroimaging results. *Psychonomic Bulletin & Review*, 18(2), 422–428. <http://dx.doi.org/10.3758/s13423-010-0048-7>.
- McCabe, D., & Castel, A. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, 107(1), 343–352. <http://dx.doi.org/10.1016/j.cognition.2007.07.017>.
- Michael, R. B., Newman, E. J., Vuorre, M., Cumming, G., & Garry, M. (2013). On the (non)persuasive power of a brain image. *Psychonomic Bulletin & Review*, 20(4), 720–725. <http://dx.doi.org/10.3758/s13423-013-0391-6>.
- Pratt, B. (2005). “Soft” science in the courtroom?: The effects of admitting neuroimaging evidence into legal proceedings. *Penn Bioethics Journal*, 1, 1–3.
- Rose, N. (2000). The biology of culpability: Pathological identity and crime control in a biological culture. *Theoretical Criminology*, 4, 5–34. <http://dx.doi.org/10.1177/136248060004001001>.
- Roskies, A. L., Schweitzer, N. J., & Saks, M. J. (2013). Neuroimages in court: less biasing than feared. *Trends in Cognitive Sciences*, 2–4. <http://dx.doi.org/10.1016/j.tics.2013.01.008>.
- Schweitzer, N. J., & Saks, M. J. (2011). Neuroimage evidence and the insanity defense. *Behavioral Sciences & the Law*, 29(4), 592–607. <http://dx.doi.org/10.1002/bsl>.
- Schweitzer, N. J., & Saks, M. J. (2012). Jurors and scientific causation: What don't they know, and what can be done about it? *Jurimetrics*, 52, 433–455.
- Schweitzer, N. J., Saks, M. J., Murphy, E. R., Roskies, A. L., Sinnott-Armstrong, W., & Gaudet, L. M. (2011). Neuroimages as evidence in a mens rea defense: No impact. *Psychology, Public Policy, and Law*, 17(3), 357–393. <http://dx.doi.org/10.1037/a0023581>.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Vincent, N. A. (2011). Neuroimaging and responsibility assessments. *Neuroethics*, 4, 35–49. <http://dx.doi.org/10.1007/s12152-008-9030-8>.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, 20(3), 470–477. <http://dx.doi.org/10.1162/jocn.2008.20040>.