# Look Again: Effects of Brain Images and Mind–Brain Dualism on Lay Evaluations of Research

**Cayce J. Hook and Martha J. Farah**

## Abstract

■ Brain scans have frequently been credited with uniquely seductive and persuasive qualities, leading to claims that fMRI research receives a disproportionate share of public attention and funding. It has been suggested that functional brain images are fascinating because they contradict dualist beliefs regarding the relationship between the body and the mind. Although previous research has indicated that brain images can increase judgments of an article's scientific reasoning, the hypotheses that brain scans make research appear more interesting, surprising, or worthy of funding have not been tested. Neither has the relation between the allure of brain imaging and dualism. In the following three studies, laypersons rated both fictional research descriptions and real science news articles accompanied by brain scans, bar charts, or photographs. Across 988 participants, we found little evidence of neuroimaging's seductive allure or of its relation to self-professed dualistic beliefs. These results, taken together with other recent null findings, suggest that brain images are less powerful than has been argued. ■



## INTRODUCTION

In their classic article entitled "Seeing Is Believing," McCabe and Castel (2008) presented evidence that scientific texts accompanied by fMRI images are perceived as more credible than texts accompanied by other graphical representations of the data. This effect was found when participants rated fictional research descriptions including intentionally flawed reasoning, as well as when participants rated a BBC article about a real research study. Across three experiments, comparing illustrations of fMRI to bar charts, a topographic map of EEG results, or no image, readers consistently rated the scientific reasoning more highly when associated with fMRI. The authors concluded that "there is something particularly persuasive about brain images with respect to conferring credibility to cognitive neuroscience data" and that "part of the scientific credibility of brain imaging as a research technique lies in the images themselves" (pp. 349–350).

Well before these findings were published, numerous authors remarked on neuroimaging's apparent hold on the public imagination: neuroimaging was variously characterized as "seductive" (Illes, De Vries, Cho, & Schraedley-Desmond, 2006, p. W27; Check, 2005, p. 254; Brammer, 2003, p. 373; Gerard & Peterson, 2003, p. 13; Merckelbach, Devilly, & Rassin, 2002, p. 492; Gordon, 2001, p. 104; Ratcliff, 1998, p. 129; Sarter, Berntson, & Cacioppo, 1996, p. 13), "dazzling" (Blakeslee, 2000, para. 1; Nicholson, 2006, para. 4; *The Lancet*, February 2004, p. 71, as cited in Lee, Broderick, & Chamberlain, 2007), and "alluring" (Khoshbin & Khoshbin, 2007, p. 182; Silbersweig & Stern, 2001, p. 1; Callicott & Weinberger, 1999, p. 120). In the years since, McCabe and Castel's (2008) article has received over 200 citations according to Google Scholar, and it has provided support for some strong denunciations of imaging research. For example, Crawford (2008) refers to neuroimaging as "that fast-acting solvent of critical faculties" (p. 65); more recently, Poole (2012) wrote that "the [fMRI] pictures, like religious icons, inspire uncritical devotion" (para. 18).

Brain imaging's alleged power to sway opinion could potentially affect more than just judgments of researchers' reasoning. According to Paul Bloom (2006), "Psychologists can be heard grousing that the only way to publish in *Science* or *Nature* is with pretty color pictures of the brain. The media, critical funding decisions, precious column inches, tenure posts, science credibility and the popular imagination have all been influenced by fMRI's seductive but deceptive grasp on our attentions." (para. 6). Weisberg (2008) worries that "the research that produces the prettiest pictures or is performed with the most expensive equipment…galvanizes public opinion, earns grants, and changes the shape of the debate" (p. 54).

As brain imaging finds new applications outside cognitive neuroscience research, its seductive allure may create new problems. Writing about the use of functional brain images in the courtroom, Khoshbin and Khoshbin (2007) caution that they are "particularly vulnerable to misuse because they are so attractive" (p. 171) and Marks (2010) warns of similar dangers in national security applications: "Given the seductive allure of neuroscientific explanations and colorful brain images, neuroscience … is

University of Pennsylvania



Journal of Cognitive Neuroscience X:Y, pp. 1–10
doi:10.1162/jocn_a_00407

particularly vulnerable to abuse" (p. 4). Johnson, Blum, and Giedd (2009) express concern over "the mistaken impression that fMRI, in particular, is an infallible mind-reading technique that can be used to establish guilt or innocence, infer 'true intentions,' detect lies, or establish competency to drive, vote, or consent to marriage." Illes et al. (2010) write of "the potential for brain scan images to create biases in the laboratory, the clinic and the courtroom" (p. 61).

Given the popularity of the "seductive allure" hypothesis concerning brain images, published replications and extensions of McCabe and Castel's studies are surprisingly scarce. To our knowledge, until this year, only two published studies further investigated the effect of brain images on ratings of scientific reasoning. Keehner, Mayberry, and Fischer (2011) examined the effects of different types of brain image on ratings of scientific credibility (compared with one another, without a nonbrain image control condition). They included five different types of functional brain images (glass brain, axial slice, 3-D brain, inflated brain, and ERP scalp topography) and assessed the effects of four image characteristics (subject-rated image complexity, realism, three dimensionality, and familiarity) on judgments of reasoning. Only three dimensionality was observed to have a significant effect, only when uncorrected for multiple comparisons, providing indirect and weak support for the hypothesis that brain images are persuasive. More recently, Gruber and Dickerson (2012) asked undergraduates to rate the credibility and quality of reasoning of a fictional science article accompanied by no image, a functional brain image, or other types of images. They failed to find an effect of image type on students' ratings of the research.

Outside the domain of research evaluation, several studies have examined the influence of brain imaging on juror decision-making. Gurley and Marcus (2008) found that participants were more likely to render a verdict of "not guilty by reason of insanity" if defendants had a prior history of psychiatric disorder or neurological damage. However, the brain images were always accompanied by additional written testimony in this study, so the influential factor may have been the testimony, and not the brain images per se. McCabe, Castel, and Rhodes (2011) examined the effects of various types of lie detection evidence on juror decision-making and found that participants rendered more guilty verdicts when fMRI evidence was described than when polygraphs, thermal face imaging, or no lie detection method were described instead (although the effect disappeared when the lie detection methods were criticized in a cross examination). Note that in this case brain images were not shown, and the information associated with the lie detection conditions differed (e.g., activation of frontal lobes for fMRI, rise in facial temperature for thermal imaging). The finding speaks directly to juror views of neuroscience evidence, but its relevance to brain images more specifically remains unclear.

The uncertainty left by McCabe and colleagues' (2011) study is addressed by two other recent studies of juror decision-making and brain images. Schweitzer and Saks (2011) and Schweitzer et al. (2011) found that, although jurors found neuroscience-based testimony more convincing than non-neuroscience-based testimony, brain images had no additional effects on juror decision-making over and above the text-based expert testimony they accompanied. Similarly, Greene and Cahill (2012) found that brain images had no greater impact than neuropsychological evidence on sentencing recommendations. Taken together, the evidence seems to suggest that expert neuroscience testimony can influence jurors, but brain images themselves play a small role, if any.

In the following studies, we attempted to conceptually replicate and expand upon the "seeing is believing" effect of brain images. Specifically, we tested claims that have been made concerning the disproportionate influence of fMRI on allocation of resources and public interest by collecting ratings of worthiness of funding, innovativeness, interest, and surprise for articles accompanied by brain scans, photographs, and bar charts.

We also tested a hypothesis concerning the reason for the public's fascination with fMRI, namely mind–brain dualism. As Bloom (2006) has put it, "we intuitively think of ourselves as nonphysical, and so it is a shock, and endlessly interesting, to see our brains at work in the act of thinking… So, when a *New York Times* article rhapsodized about neural correlates of passion ('Watching New Love As It Sears the Brain'), the interest of the article for the average reader did not lie in the details about the role of the caudate nucleus. Rather, it lay in the fact that the brain is involved at all in anything as interesting and personal as falling in love" (para. 8). According to a *New York Times* article on fMRI, the "the sheer delight, the true amazement" people feel when viewing brain images is attributable to their surprise "that there is a measurable physical response in the brain" to love, jealousy, and other seemingly "magic" phenomena (Lucy Brown, quoted in Carey, 2006, para. 11). Given that beliefs about the relationship between the mind and brain vary from person to person, we hypothesized that brain images would elicit different responses from people who express dualist versus physicalist beliefs.

## EXPERIMENT 1

### Methods

#### *Participants*

Participants were recruited from Amazon Mechanical Turk (MTurk) and were paid 50 cents for their participation, a typical compensation rate for the site. MTurk samples tend to be younger, more female, and more highly educated than the general U.S. population, as well as reporting lower annual income (Paolacci, Chandler, & Ipeirotis, 2010). Nonetheless, studies indicate that MTurk samples are at least as diverse as traditional university and internet samples, and the data obtained from these subject pools are comparable (e.g., Mason & Suri, 2012; Buhrmester, Kwang, & Gosling, 2011). For example, a well-known study by Weisberg, Keil, Goodstein, Rawson, and Gray (2008), pertaining to the effects of neuroscience information on judgments of reasoning, was recently replicated using an MTurk sample (Michael, Newman, Vuorre, Cumming, & Garry, 2013).

In an effort to minimize comprehension problems, we recruited only U.S. residents. The study included 180 participants (124 women, 56 men; median age = 31 years, range = 18–70 years). One hundred seventy-five (97.2%) were native English speakers; 63% reported holding a college degree or higher, 30% reported having completed some college, and 7% reported having completed high school or less. Gender, age, and education levels were thus biased in ways characteristic of the MTurk population. Although these biases limit the generalizability of our results; this sample is nevertheless more diverse than the undergraduate student samples used in previous studies (e.g., Keehner et al., 2011; McCabe & Castel, 2008).

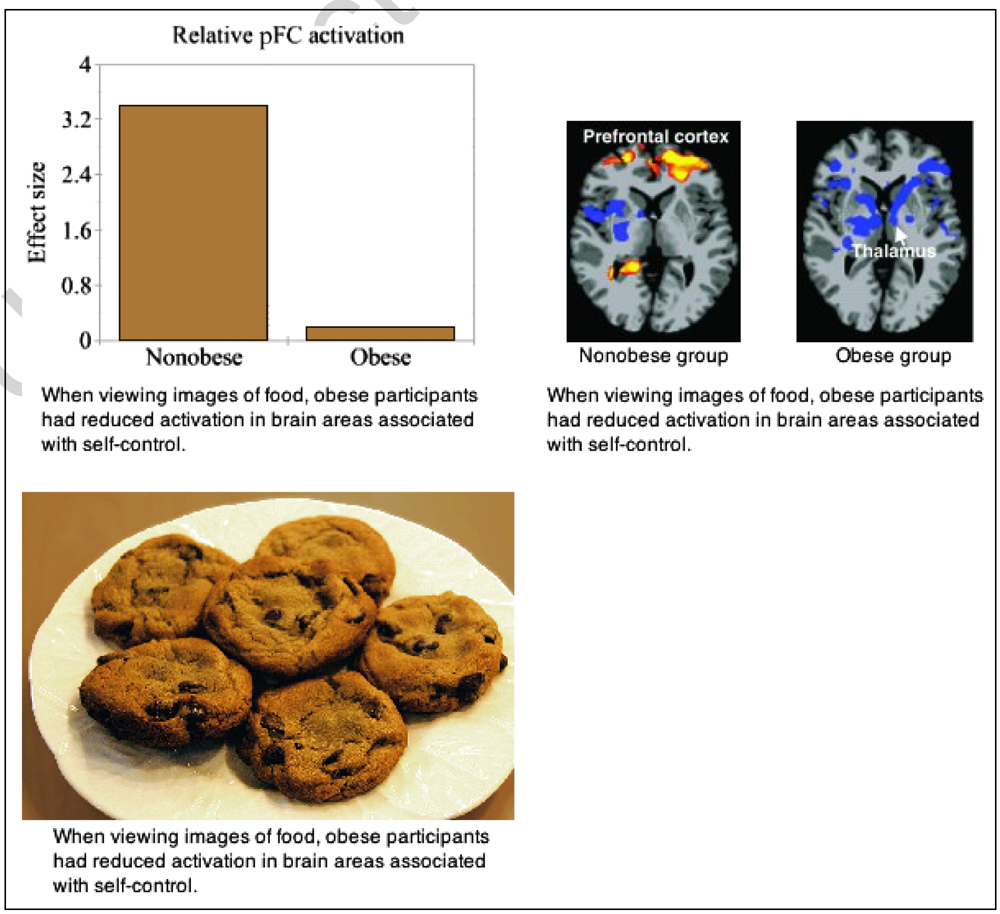
### *Materials*

Three four- to five-sentence fictional research descriptions were created: "Sleep and Memory," "Videogames and Attention," and "Obesity and Self Control." Appropriate fMRI images and photographs were selected from academic journals and stock photo websites, and bar charts were created to represent the fictional results. All texts described plausible research experiments, and all images bore identical captions (e.g., Figure 1).

For each research description, there were five rating scales, presented as 100-point slider scales with 0 labeled as "strongly disagree" and 100 as "strongly agree." Ratings were elicited for reasoning, as in McCabe and Castel's (2008) study, with the statement "The scientific reasoning in the article made sense." In addition, four other qualities were rated: interest ("This research is interesting"), surprise ("This study's finding is surprising"), innovation ("This research is innovative"), and worthiness of funding ("This research is worthy of taxpayer funding").

Our dualism scale was composed of items from two published dualism scales, selected for being as easy to understand as possible (in the authors' judgments). From the four-statement survey described by Demertzi et al. (2009), we used the following three items: "The mind is fundamentally physical" (reverse coded), "Some spiritual part of us survives after death," and "Each of us has a soul that is separate from the body." We excluded the item "The mind and brain are two separate things," because we felt it might be subject to multiple interpretations. We also included four statements from Stanovich's (1989) 27-item dualism scale, selected for being relatively unambiguous yet describing dualism in different

**Figure 1.** Sample stimuli from Experiment 1.



ways: "All mental processes are the result of activity in the nervous system" (reverse coded), "When I use the word 'mind,' it is just a shorthand term for the complicated things my brain does" (reverse coded), "The mind is a nonmaterial substance that interacts with the brain to determine behavior," and "Minds are inside brains but not the same as brains."

A brief demographic questionnaire elicited gender, age, native language (text box), race/ethnicity (text box), level of education (choices were: Less than high school, High school diploma, Some college, College degree, Some graduate school, Master's degree, or Doctorate or professional degree: Ph.D., M.D., J.D.), religious background (text box). Finally, an assessment of theism was phrased as "Please indicate the extent to which you believe in the existence of a Supreme Being or God (or supreme beings/Gods)" and answered with a 100-point slider scale from "not at all" to "very strongly."

### *Procedure*

Each participant was randomly assigned to one of the following three image conditions: fMRI scans, stock photographs, or bar charts. That is, each participant read and evaluated the articles accompanied by a single type of image.

Participants were informed that the investigators were interested in understanding people's attitudes toward different types of research. They first responded to the seven-item dualism questionnaire. One of the three research descriptions was then presented at random. To verify that participants paid sufficient attention to the task, after participants finished reading the text, another screen asked participants to provide a brief summary of what they had read. Participants were again presented with the text and image (fMRI, photo, or bar chart) and asked to rate the research on the five dimensions listed above. This process was then repeated with the remaining two texts. After they had read and rated all three research descriptions, participants responded to the demographic information questions.

## Results and Discussion

Before examining participants' ratings of the articles, the first author reviewed participants' responses to the article summary questions for evidence of attention to the task. Although the responses differed in their level of detail, all contained a minimum of one complete sentence describing the study's conclusion.

Responses to the dualism scale were analyzed for reliability. Cronbach's alpha for the scale from the current sample was .817, and all item–test correlations were between .4 and .7. Responses to the seven dualism items were then averaged to create a single dualism score for each participant (with scores close to 0 reflecting physicalism and scores close to 100 reflecting dualism). Dualism score was significantly correlated with strength of belief in a supreme being, god or gods ($r = .613$, $p < .001$), but not with educational attainment ($r = -0.059$, $p = .429$). Strong physicalist views were more common than strong dualist views, with 21 participants scoring below 15 (strongly physicalist) but only one participant scoring above 85 (strongly dualist). One third of the sample scored between 40 and 60, indicating relatively equal agreement with both physicalist and dualist statements. Rather than splitting participants along the median and filling each of two groups with participants having similar dualism scores, we classified participants scoring in the lowest third as "physicalists" (score range = 0–34), those in the top third as "dualists" (score range = 53.29–85.71), and those in the middle third as "intermediate" (score range = 34.71–53).

Ratings for "interesting," "surprising," "innovative," "reasoning," and "worthy of funding" were averaged across the three research descriptions to create five dependent variables for each participant.

A two-way MANOVA (image condition, dualism group) revealed a statistically significant difference in ratings of research among the three image conditions, $F(10, 334) = 2.635$, $p < .005$; Wilk's $\lambda = 0.859$, partial $\eta^2 = .073$. There was also a significant effect of dualism, $F(10, 334) = 2.399$, $p < .01$; Wilk's $\lambda = 0.87$, partial $\eta^2 = .067$. However, there was no interaction between image condition and dualism, $F(20, 662) = .437$, $p = .986$; Wilk's $\lambda = 0.949$, partial $\eta^2 = .013$), indicating that the overall effect of image type on ratings was no different for physicalists and dualists.

A priori pairwise multivariate contrasts indicated significant differences between the fMRI and photo conditions, $F(5, 113) = 2.8$, $p = .02$, Hotelling's $T^2 = 14.76$, as well as the fMRI and chart conditions, $F(5, 111) = 3.81$, $p = .003$, Hotelling's $T^2 = 19.78$, and no significant difference between the photo and chart conditions, $F(5, 118) = 1.94$, $p > .09$, Hotelling's $T^2 = 10.00$, suggesting that the overall evaluation of research descriptions was more positive in the presence of an fMRI image than either a photo or bar chart, when all dependent measures are considered together.

Means for each dimension of research evaluation by each image condition and dualism group are shown in Table 1. To test the "seeing is believing" hypothesis, we used a two-way ANOVA to assess the effects of image type on judgments of reasoning. This analysis yielded no significant effect of image condition, dualism, or their interaction on participants' ratings of scientific reasoning, $F(2, 171) = 1.67$, $p = .19$. Additional ANOVAs were carried out to determine whether an effect of image condition, dualism, or their interaction held for any of the other dimensions of research evaluation. There were significant effects of image condition on "interesting," $F(2, 171) = 5.05$, $p < .01$, "innovative," $F(2, 171) = 6.29$, $p < .005$, and "worthy of funding," $F(2, 171) = 4.81$, $p < .01$, but not for "surprising," $F(2, 171) = 2.34$, $p = .099$. Dualism group had a

**Table 1.**

| | *Interesting* | *Surprising* | *Innovative* | *Reasoning* | *Funding* |
|---|---|---|---|---|---|
| fMRI, $n = 56$ | $M = 78.05$, $SD = 14.82$ | $M = 39.03$, $SD = 16.04$ | $M = 54.92$, $SD = 18.48$ | $M = 77.61$, $SD = 18.38$ | $M = 53.70$, $SD = 23.75$ |
| Photo, $n = 63$ | $M = 70.37$, $SD = 15.81$ | $M = 41.63$, $SD = 18.25$ | $M = 49.13$, $SD = 21.17$ | $M = 71.80$, $SD = 17.60$ | $M = 45.95$, $SD = 24.29$ |
| Chart, $n = 61$ | $M = 69.70$, $SD = 15.72$ | $M = 34.70$, $SD = 15.49$ | $M = 42.42$, $SD = 16.71$ | $M = 74.64$, $SD = 16.19$ | $M = 40.81$, $SD = 23.65$ |
| Dualists, $n = 60$ | $M = 72.78$, $SD = 15.41$ | $M = 37.78$, $SD = 18.50$ | $M = 49.06$, $SD = 18.76$ | $M = 76.21$, $SD = 17.46$ | $M = 45.15$, $SD = 23.33$ |
| Intermediate, $n = 60$ | $M = 73.89$, $SD = 15.24$ | $M = 41.46$, $SD = 16.24$ | $M = 49.76$, $SD = 20.14$ | $M = 73.31$, $SD = 15.46$ | $M = 40.85$, $SD = 22.99$ |
| Physicalists, $n = 60$ | $M = 70.91$, $SD = 16.93$ | $M = 36.17$, $SD = 15.46$ | $M = 47.15$, $SD = 19.75$ | $M = 74.19$, $SD = 19.38$ | $M = 53.86$, $SD = 25.21$ |

significant effect on ratings of funding-worthiness, $F(2, 171) = 5.45$, $p < .01$, but not on "surprising" or any of the other dependent variables (all $p$s $> .2$), and there were no significant interactions between image condition and dualism group (all $p$s $> .1$).

Planned comparisons were carried out to determine whether fMRI improved ratings relative to the other two image conditions. This was the case for ratings of interestingness: texts with fMRI images were rated as more interesting than those with photographs or bar charts, $t = 2.73$, $p < .01$ and $t = 2.81$, $p < .01$, respectively. Additionally, the fMRI condition was rated as more innovative, $t = 3.54$, $p = .001$, and worthy of funding, $t = 3.1$, $p < .005$, than the bar chart condition, but not the photo condition (both $p > .8$).

Differences between dualism groups were tested using Tukey's post hoc comparisons. The only difference that was statistically significant was physicalists' higher rating of funding worthiness compared with intermediate participants ($p < .01$).

The results of Experiment 1 support the hypothesis that fMRI images add appeal to short descriptions of research. However, we did not find an effect of image type on judgments of scientific reasoning. Compared with photos, fMRI improved the rating of only one dimension of five: interestingness. The hypothesis that dualism is partly responsible for the appeal of brain images was also not supported. Although our measure of dualism had good internal consistency and showed a strong positive relation to theism, it did not predict the effect of brain images on evaluations of research.

## EXPERIMENT 2

The second experiment was a new attempt to test the "seeing is believing" effect and its moderation by dualism. To increase our ability to find these effects, we made the following changes. In light of the statements about brain imaging and dualism quoted above, one might expect brain images to be more striking, fascinating, and persuasive when they illustrate the neural bases of more emotional and personal aspects the mind; we therefore substituted real science news stories relevant to temptation, love, pain, and bullying for the more emotionally neutral topics used in Experiment 1. In addition, we changed our measure of dualism to capture different and possibly more relevant aspects of dualism. We also increased the sample size and dropped the bar chart condition, resulting in almost twice as many participants in each of the brain image and photo conditions.

### Methods

*Participants*

Two hundred U.S. residents initially participated. Six participants were excluded after their responses indicated a failure to read the articles, leaving 194 participants (106 men, 87 women, 1 unreported; median age = 29 years; range = 17–80 years). As in Experiment 1, participants were recruited from Amazon Mechanical Turk and were paid 50 cents for their participation. The majority were native English speakers (97.9%); 53% had completed college or higher, 36% had attended some college, 10% had a high school diploma, one had not completed high school, and one did not report an education level.

### *Materials and Procedure*

Three articles, "Obese May Be Less Able to Control Food Impulses," "Pain May Be Pleasurable for Some Bullies," and "Baby Love May Be Hard-Wired in Human Brain," were selected from on-line news sources (Reuters, ABC News, and MSNBC, respectively) and were edited to a length of approximately 400 words. fMRI images were retrieved from the journal articles described in each news piece, and complementary photographs were selected from stock photo websites (Figure 2).

Participants responded to the same statements as in Experiment 1, with the following modifications. The funding question was changed to "This research is worthy of taxpayer funding (e.g., through organizations such as the National Institutes of Health)." In addition, to more closely match the outcome measure used in McCabe and Castel's (2008) Experiment 3, which examined the effect of fMRI images on judgments of a real news article, we changed the wording of the "scientific reasoning" question to the following: "Do you agree or disagree with the article's conclusion that [obese people may be less able to shut off the parts of the brain that drive food cravings?]/[young people with conduct disorder enjoy seeing others in pain and lack the ability to control potentially inappropriate emotions?]/[the lure of babies stems from specific brain circuits?]"

We modified the dualism scale to include the following five items: "Though our bodies die, our minds (consciousness, memory, will) can survive" (adapted from Bloom, 2006); "Some nonmaterial part of me (my mind, soul, or spirit) determines my behavior"; "All of my conscious experience is the result of activity in my nervous system" (reverse coded); "My mind (consciousness, memory, will) is an emergent property of my brain and cannot be separated from it" (reverse coded); and "The mind and the brain are the same thing. When I use the word 'mind,' it is just a shorthand term for the things my brain does" (reverse coded; adapted from Stanovich, 1989).

The procedure of Experiment 2 was similar to that of Experiment 1, with two changes: participants rated each article on a 5-point scale, and the modified dualism questionnaire was presented randomly either before or after the three articles.

## Results and Discussion

Participants' responses to the article summary questions were reviewed by the first author. Participants whose responses consisted of one or two words (such as "good" or "more brains"), random keystrokes (such as "kl;a; kldd"), or statements irrelevant to the articles (e.g., "There are alot of people with bad health problems") were excluded from analysis ($n = 6$), leaving 194 participants.

Ratings of the three articles were averaged to create five dependent variables per participant. Responses to the dualism scale were confirmed to be reliable (Cronbach's alpha = .828; all item-scale correlations between .56 and .7) and were then averaged to create a single dualism score for each participant. As in Experiment 1, dualism was significantly correlated with theism ($r = .54$, $p < .001$), but not with educational attainment ($r = -.013$, $p > .8$). Participants were classified by thirds as before: "physicalists" (score range = 1.0–2.2), "dualists" (score range = 3.0–5.0), and "intermediate" (score range = 2.2–3.0).

Means for each dimension of research evaluation by image condition and dualism group are shown in Table 2.



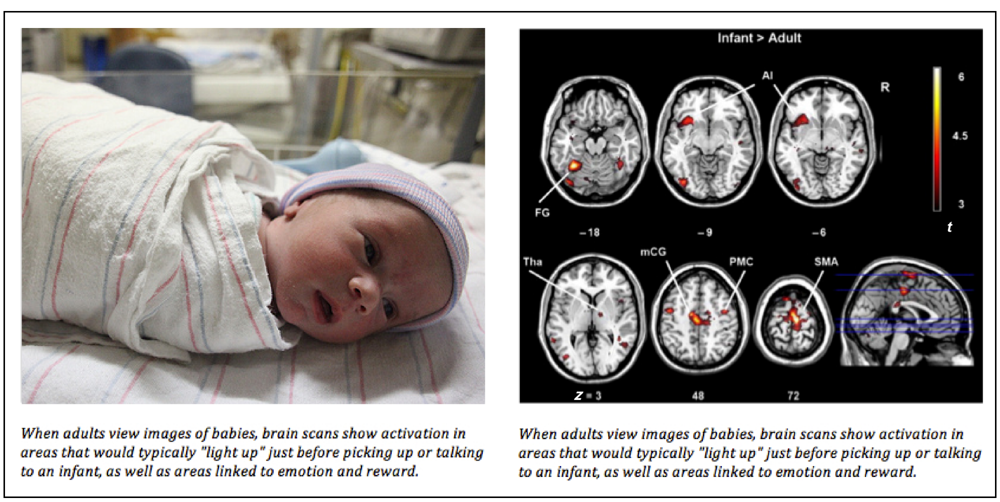



**Figure 2.** Sample images from Experiment 2.

**Table 2.**

| | *Interesting* | *Surprising* | *Innovative* | *Worthy of Funding* | *Agree with Conclusion* |
|---|---|---|---|---|---|
| fMRI, $n$ = 98 | $M$ = 3.60, $SD$ = .70 | $M$ = 2.47, $SD$ = .89 | $M$ = 3.22, $SD$ = .80 | $M$ = 3.03, $SD$ = .91 | $M$ = 3.75, $SD$ = .64 |
| Photo, $n$ = 96 | $M$ = 3.46, $SD$ = .76 | $M$ = 2.30, $SD$ = .72 | $M$ = 3.04, $SD$ = .74 | $M$ = 2.85, $SD$ = 1.10 | $M$ = 3.80, $SD$ = .61 |
| Dualists, $n$ = 65 | $M$ = 3.60, $SD$ = .72 | $M$ = 2.45, $SD$ = .77 | $M$ = 3.16, $SD$ = .75 | $M$ = 2.80, $SD$ = 1.09 | $M$ = 3.57, $SD$ = .69 |
| Intermediate, $n$ = 64 | $M$ = 3.48, $SD$ = .72 | $M$ = 2.38, $SD$ = .86 | $M$ = 3.07, $SD$ = .77 | $M$ = 2.91, $SD$ = .97 | $M$ = 3.89, $SD$ = .58 |
| Physicalists, $n$ = 65 | $M$ = 3.51, $SD$ = .76 | $M$ = 2.33, $SD$ = .82 | $M$ = 3.17, $SD$ = .82 | $M$ = 3.11, $SD$ = .94 | $M$ = 3.86, $SD$ = .56 |



In contrast to the results of the previous experiment, a two-way MANOVA revealed no effect of image type, $F(5, 184) = .958, p > .4$ (Wilk's $\lambda = 0.975$, partial $\varepsilon^2 = .025$), on overall evaluation of research. There was a significant effect of dualism group on overall ratings, $F(10, 368) = 2.092, p < .05$ (Wilk's $\lambda = .895$, partial $\varepsilon^2 = .054$), but no interaction between image condition and dualism group, $F(10, 368) = .672, p > .7$ (Wilk's $\lambda = .964$, partial $\varepsilon^2 = .018$).

Two-way ANOVAs revealed a significant effect of dualism group on "agreement with conclusion," $F(2, 188) = 5.082, p < .01$; Tukey's post hoc comparisons indicated that dualists expressed significantly lower levels of agreement with the conclusions of the articles than both physicalists ($p = .023$) and intermediates ($p = .013$). There were no other significant effects (all $p$s > .09).

In contrast to the first experiment, in this larger experiment, we did not find an effect of fMRI on laypersons' overall evaluation of research nor their judgments of any one of the five dimensions rated. Dualists were no more likely to be influenced by fMRI than physicalists (although overall they were less convinced of the stated research conclusions than physicalists).

In summary, whereas the first two studies are consistent regarding the lack of interaction between dualism and image condition, they partially diverge on the effect of fMRI itself. Neither experiment replicated the "seeing is believing" effect on judgments of scientific reasoning, but the first experiment demonstrated an effect of image condition overall and enhanced evaluation of "interestingness" with fMRI compared with photos, whereas the second experiment did not. To better understand the effects of fMRI images on the evaluation of research, we undertook a third experiment, with an even larger sample and hence increased power.

## EXPERIMENT 3

### Methods

#### *Participants*

Six hundred sixty-eight U.S. residents initially participated. Participants were recruited from Amazon Mechanical Turk and were paid 50 cents for their participation. MTurk users who completed Study 2 were prevented from participating in Study 3. Fifty-four participants' responses were judged inadequate according to the criteria listed above; these data were excluded from analysis, leaving 614 participants (374 women, 239 men, 1 unreported; median age = 29 years, range = 18–74 years). Native language and educational attainment were similar to those in Experiment 2.

#### *Materials and Procedure*

The materials were identical to those in Experiment 2, with two exceptions: A bar chart condition was added, and participants were only asked the questions pertaining to interestingness, worthiness of taxpayer funding, and agreement with conclusion. We chose to focus on these outcomes because interestingness showed the strongest effects in Experiment 1 and because credibility and funding worthiness have been most discussed in the previous literature. Display order of the dualism scale (before or after reading the three articles) was recorded and analyzed.

### Results and Discussion

Table 3 shows the means and standard deviations for each dimension of research evaluation by image condition and dualism group. As in Experiment 2, two-way MANOVA revealed a significant effect of dualism group on ratings of research, $F(6, 1206) = 6.25$, $p < .0005$ (Wilk's $\lambda = .94$, partial $\eta^2 = .03$), no effect of image condition, $F(6, 1206) = .826$, $p > .5$ (Wilk's $\lambda = .992$, partial $\eta^2 = .004$), and no interaction, $F(12, 1595.68) = 1.23$, $p > .2$ (Wilk's $\lambda = .976$, partial $\eta^2 = .008$).

Two-way ANOVAs were carried out to examine the effects of image type and dualism group on the three dependent variables separately. Again, image type had no significant effect on judgments of credibility or any other ratings of research, all *F*s $\leq 1$, all *p*s $> .3$.

There were significant effects of dualism on ratings of funding worthiness and agreement with conclusion [$F(2, 605) = 8.885$, $p < .0005$ and $F(2, 605) = 5.39$, $p = .005$, respectively] and no interactions (all *F*s $< 1.5$, all *p*s $>$ .2). Tukey's post hoc analyses revealed that regardless of image type, physicalists rated research as significantly more worthy of funding than dualists and intermediates ($p < .0005$; $p < .05$) and agreed more strongly with the articles' conclusions than dualists ($p < .005$).

There was a small but statistically significant effect of the order in which the dualism scale was presented on participants' dualism scores, such that participants who answered the dualism scale before reading the three articles tended to agree more with dualist statements than those who answered the questions after reading the three articles ($t = 2.21$, $p < .05$; means = 2.84 and 2.69, respectively). This is consistent with the articles' focus on cognitive neuroscience priming physicalist beliefs. As in the previous two experiments, dualism scores were significantly correlated with religiosity ($r = .57$, $p < .001$) but not with educational attainment ($r = -.04$, $p > .2$).

## GENERAL DISCUSSION

It is widely believed that brain images are overly influential. Much to our surprise, our findings did not support this view. Across three experiments with a total of 988 participants, we found no evidence for the "seeing is believing" effect of brain scans. One of the experiments showed that fMRI enhanced laypersons' overall evaluations of research, but this effect was absent in the two larger experiments, and in no study was the research in question viewed as more credible when accompanied by a brain image.

Under these circumstances, the question of whether dualism is in part responsible for the allure of brain imaging is largely moot, and indeed we failed to find differences in the effect of imaging as a function of dualist belief. An unexpected finding was the trend across experiments for physicalists to be more in favor of funding the research described and more inclined to agree with the reasoning or conclusions.

It is worth noting that we only assessed overt dualism; the possibility remains that covert dualistic intuitions could influence perceptions of brain images. However, in that case, we might expect articles accompanied by brain images to be judged as more surprising than those without; this hypothesis was also not confirmed.

The present results do not address the broader question of whether neuroscience in general is overly influential. Weisberg et al. (2008) argued that it may be, at least when poor explanations of psychological phenomena are being evaluated. Their research demonstrated that poor explanations were judged more satisfying when they included references to neuroscience, even when the neuroscience information was superfluous. On the other hand, adding irrelevant neuroscience information to "good" explanations had no significant effect on laypersons' judgments.

In contrast to the stimuli used by Weisberg et al. (2008) and some of those used by McCabe and Castel



**Table 3.**

| | *Interesting* | *Worthy of Funding* | *Agree with Conclusion* |
|---|---|---|---|
| fMRI, $n = 198$ | $M = 3.65$, $SD = .64$ | $M = 3.12$, $SD = .89$ | $M = 3.72$, $SD = .62$ |
| Photo, $n = 203$ | $M = 3.61$, $SD = .77$ | $M = 3.13$, $SD = .90$ | $M = 3.74$, $SD = .68$ |
| Chart, $n = 213$ | $M = 3.64$, $SD = .65$ | $M = 3.06$, $SD = .94$ | $M = 3.67$, $SD = .64$ |
| Dualists, $n = 205$ | $M = 3.66$, $SD = .72$ | $M = 2.94$, $SD = .89$ | $M = 3.61$, $SD = .69$ |
| Intermediate, $n = 204$ | $M = 3.61$, $SD = .71$ | $M = 3.07$, $SD = .86$ | $M = 3.69$, $SD = .56$ |
| Physicalists, $n = 205$ | $M = 3.63$, $SD = .63$ | $M = 3.3$, $SD = .94$ | $M = 3.82$, $SD = .66$ |

(2008), none of the texts we used contained reasoning errors. Our finding that brain images did not influence ratings of credibility is thus not inconsistent with the findings reported by Weisberg et al. (2008). However, our findings contradict those of McCabe and Castel's (2008) Experiment 3, in which the addition of a brain image improved judgments of credibility for a real news article that contained no logical errors.

Other recent replication attempts cast further doubt on the persuasive power of brain images. Michael et al. (2013) conducted 10 direct replications of McCabe and Castel's (2008) Experiment 3. A meta-analysis including these and McCabe and Castel's original results found that brain images exerted "little to no influence" on judgments of credibility. Schweitzer, Baker, and Risko (submitted) report three experiments, including one near-direct replication, none of which found an effect of brain images on judgment.

Has the public simply grown more savvy about neuroscience in recent years? That the findings of Weisberg et al. (2008) can still be successfully replicated suggests that this is not the case (Michael et al., 2013). Instead, it seems that neuroscience evidence may enhance the credibility of poorly reasoned arguments, even if brain images per se do not. This interpretation is consistent with several recent studies on juror decision-making, discussed above, which found that, although neuroscience-based testimony can be persuasive, brain images themselves exert little additional influence. Taken together with these findings, our results paint a clear picture: When it comes to brain images, seeing is not necessarily believing.

## UNCITED REFERENCE

Farah & Hook, 2013

### Acknowledgments

We thank Robert Michael, Maryanne Garry, and Nick Schweitzer for sharing their findings and ideas with us. We also thank our colleagues at the Center for Neuroscience and Society and the Intuitive Neuroscience Working Group, as well as two anonymous reviewers, for helpful discussion. The writing of this article was supported by NIH grant R01 HD055689 and NSF grant 0933919.

Reprint requests should be sent to Cayce J. Hook, Center for Neuroscience and Society, University of Pennsylvania, 3720 Walnut St., Philadelphia, PA 19104, or via e-mail: caycehook@gmail.com, hook@sas.upenn.edu.

## REFERENCES

Blakeslee, S. (2000, March 14). Just what's going on inside that head of yours? *The New York Times*. Retrieved from www.nytimes.com/2000/03/14/health/just-what-s-going-on-inside-that-head-of-yours.html.

Bloom, P. (2006, June 26). Seduced by the flickering lights of the brain. *Seed Magazine*. Retrieved from seedmagazine.com/content/article/seduced_by_the_flickering_lights_of_the_brain/.

Brammer, M. (2003). Editorial. *Statistical Methods in Medical Research, 12,* 373–374.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5.

Callicott, J. H., & Weinberger, D. R. (1999). Functional brain imaging: Future prospects for clinical practice. In S. H. Weissman, M. Sabshin, & H. Eist (Eds.), *Psychiatry in the new millennium*. Arlington, VA: American Psychiatric Publishing.

Carey, B. (2006, February 5). Searching for the person in the brain. *The New York Times*. Retrieved from www.nytimes.com/2006/02/05/weekinreview/05carey.html.

Check, E. (2005). Ethicists urge caution over emotive power of brain scans. *Nature, 435,* 254.

Crawford, M. B. (2008). The limits of neuro-talk. *The New Atlantis, 19,* 65–78.

Demertzi, A., Liew, C., Ledoux, D., Bruno, M.-A., Sharpe, M., Laureys, S., et al. (2009). Dualism persists in the science of mind. *Annals of the New York Academy of Sciences, 1157,* 1–9.

Farah, M. J., & Hook, C. J. (2013). The seductive allure of "seductive allure." *Perspectives on Psychological Science, 8,* 88–90.

Gerard, E., & Peterson, B. S. (2003). Developmental processes and brain imaging studies in Tourette syndrome. *Journal of Psychosomatic Research, 55,* 13–22.

Gordon, E. (2001). Integrative psychophysiology. *International Journal of Psychophysiology, 42,* 95–108.

Greene, E., & Cahill, B. S. (2012). Effects of neuroimaging evidence on mock juror decision making. *Behavioral Sciences & the Law, 30,* 280–296.

Gruber, D., & Dickerson, J. A. (2012). Persuasive images in popular science: Testing judgments of scientific reasoning and credibility. *Public Understanding of Science, 21,* 938–948.

Gurley, J. R., & Marcus, D. K. (2008). The effects of neuroimaging and brain injury on insanity defenses. *Behavioral Sciences & the Law, 26,* 85–97.

Illes, J., De Vries, R., Cho, M. K., & Schraedley-Desmond, P. (2006). ELSI priorities for brain imaging. *The American Journal of Bioethics, 6,* W24–W31.

Illes, J., Moser, M. A., McCormick, J. B., Racine, E., Blakeslee, S., Caplan, A., et al. (2010). Neurotalk: Improving the communication of neuroscience research. *Nature Reviews Neuroscience, 11,* 61–69.

Johnson, S. B., Blum, R. W., & Giedd, J. N. (2009). Adolescent maturity and the brain: The promise and pitfalls of neuroscience research in adolescent health policy. *The Journal of Adolescent Health, 45,* 216–221.

Keehner, M., Mayberry, L., & Fischer, M. (2011). Different clues from different views: The role of image format in public perceptions of neuroimaging results. *Psychonomic Bulletin & Review, 18,* 422–428.

Khoshbin, L. S., & Khoshbin, S. (2007). Imaging the mind, minding the image: An historical introduction to brain imaging and the law. *American Journal of Law & Medicine, 33,* 171.

Lee, N., Broderick, A. J., & Chamberlain, L. (2007). What is "neuromarketing"? A discussion and agenda for future research. *International Journal of Psychophysiology, 63,* 199–204.

Marks, J. H. (2010). A neuroskeptic's guide to neuroethics and national security. *AJOB Neuroscience, 1,* 4–12.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44,* 1–23.

McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107,* 343–352.

McCabe, D. P., Castel, A. D., & Rhodes, M. G. (2011). The influence of fMRI lie detection evidence on juror decision-making. *Behavioral Sciences & the Law, 29,* 566–577.

Merckelbach, H., Devilly, G. J., & Rassin, E. (2002). Alters in dissociative identity disorder: Metaphors or genuine entities? *Clinical Psychology Review, 22,* 481–497.

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*. Advance on-line publication. doi:10.3758/s13423-013-0391-6

Nicholson, C. (2006). Thinking it over: fMRI and psychological science. *APS Observer, 19,* 20–25.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon's Mechanical Turk. *Judgment and Decision-making, 5,* 411–419.

Poole, S. (2012, September 6). Your brain on pseudoscience: The rise of popular neurobollocks. *The New Statesman*. Retrieved from www.newstatesman.com/culture/books/2012/09/your-brain-pseudoscience.

Ratcliff, R. (1998). The role of mathematical psychology in experimental psychology. *Australian Journal of Psychology, 50,* 129–130.

Sarter, M., Berntson, G. G., & Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience: Toward strong inference in attributing function to structure. *American Psychologist, 51,* 13–21.

Schweitzer, N. J., Baker, D. A., & Risko, E. F. (submitted). Fooled by the brain: Re-examining the influence of neuroimages.

Schweitzer, N. J., & Saks, M. J. (2011). Neuroimage evidence and the insanity defense. *Behavioral Sciences & the Law, 29,* 592–607.

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,* 357–393.

Silbersweig, D. A., & Stern, E. (2001). Functional neuroimaging and neuropsychology: Convergence, advances and new directions. *Journal of Clinical and Experimental Neuropsychology, 23,* 1–2.

Stanovich, K. E. (1989). Implicit philosophies of mind: The dualism scale and its relation to religiosity and belief in extrasensory perception. *Journal of Psychology, 123,* 5.

Weisberg, D. S. (2008). Caveat lector: The presentation of neuroscience information in the popular media. *Scientific Review of Mental Health Practice, 6,* 51–56.

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,* 470–477.

# AUTHOR QUERIES

**AUTHOR PLEASE ANSWER ALL QUERIES**

During the preparation of your manuscript, the questions listed below arose. Kindly supply the necessary information.

1. Please provide page numbers of Callicott & Weinberger, 1999.
2. Please update publication status/details of Michael et al., 2013; Schweitzer et al., submitted.
3. Uncited reference: This section comprises reference included in the reference list but without any matching entry in the text. Please position in the text where appropriate or, alternatively, delete this item from the reference list.
4. Please provide captions for Tables 1–3.
5. Figures 1 and 2 with pixelated text, lines and images. Please provide replacement figures when necessary.

**END OF ALL QUERIES**