# Lay Theories of the Mind/Brain Relationship and the Allure of Neuroscience

**Diego Fernandez-Duque**

> *Allure: the quality of being powerfully and mysteriously attractive or fascinating.*

The '90s were declared the Decade of the Brain by United States' President George H. W. Bush, who at the time of his proclamation reflected that "the human brain, a 3-pound mass of interwoven nerve cells that controls our activity, is one of the most magnificent and mysterious wonders of creation. The seat of human intelligence, interpreter of senses, and controller of movement, this incredible organ continues to intrigue scientists and layman alike" (Library of Congress' Website, https://www.loc.gov/loc/brain/). Upon reading those sentences, it is hard not to feel some amount of sympathy for the author, as he struggles to find the words to best describe the relation between brain and mind. Absent from the proclamation are the attributes we most often think of as uniquely human: a notion of self, rational thought, language, free will. Similarly, no reference is made about cognition, consciousness, emotion, or the mind. Instead, the president plays it safe, mentions *intelligence*, focuses on the brain's ability for sensation decoding and movement control, and calls it a day.

In all fairness, most people probably could not have done much better. Philosophers have been debating the mind/brain question for centuries. In the 1600s, Descartes proposed that humans were a combination of body and mind (Descartes, 1984/1641). For Descartes, the body was part of the natural world and as such it was bound by the laws of nature. But the mind, Descartes thought, was capable of abilities that were uniquely human, including moral evaluation, appreciation of beauty, and free will. Philosophers have come a long way since the days of Descartes, and his *substance dualism* has run out of favor in most philosophical circles. However, it is unclear where ordinary folk stand on the issue. Many religious beliefs that are popular across the globe, such as beliefs in the afterlife and in

D. Fernandez-Duque (✉)
Psychology Department, Villanova University,
800 Lancaster Ave, Villanova, PA 19085, USA
e-mail: diego.fernandezduque@villanova.edu

© Springer International Publishing AG 2017

C.M. Zedelius et al. (eds.), *The Science of Lay Theories*,
DOI 10.1007/978-3-319-57306-9_9

diego.fernandezduque@villanova.edu

the existence of the soul, necessitate a dualist concept of mind/brain relation (Bering & Bjorklund, 2004). Most people believe both in free will and moral responsibility (Monroe & Malle, 2010; Nahmias, Morris, Nadelhoffer, & Turner, 2005), holding others responsible for actions only in situations where a choice to act differently was available; the dualist view can easily accommodate such a perspective. On the other hand, the current explosion of research in neuroscience, with descriptions in the popular media of brains doing things that minds were supposed to do (Racine, Bar-Ilan, & Illes, 2005), has perhaps begun to challenge the beliefs that ordinary people have inherited from Descartes. For researchers interested in studying people's *common sense* theories about brain and mind, the challenge ahead is to find methods that faithfully capture those lay theories.

## How to Study Lay Theory of Mind/Brain Relation?

Several approaches exist to probe lay theories of mind/brain, each of which has its own sets of strengths and weaknesses. One option is to ask people directly about their beliefs. This is sometimes done in the form of questionnaires probing different variants of dualism (e.g., mind and body are qualitatively distinct), materialism (mind and body are the same or fundamentally united), emergentism (mind and brains are qualitatively different but interdependent), or some other form of—*isms* (Demertzi et al., 2009; Stanovich, 1989). Questionnaires allow researchers to explore possible associations with religious beliefs, and with other folk beliefs like belief in the afterlife and belief in free will. Other times, researchers ask simple questions—such as "do you need a brain to think?"—in order to assess people's common sense beliefs. An obvious strength of this approach is its face validity. If we want to know what people think about X, asking them directly what they think about X seems a sensible first step. But this approach risks running into problems if people are inconsistent, both in the sense that their answers might differ depending on how they are probed, and in the sense that their reflective answer might not align with the set of background beliefs they regularly hold for judgments in everyday life. Answers to questionnaires may also be susceptible to educational biases and cultural influences. Depending on the research question, such variability may be seen as a strength or a weakness.

To overcome these limitations, lay theory researchers sometimes probe participants' knowledge in more indirect ways. Instead of asking people to reflect on the nature of the mind/brain relation—a prospect overloaded with philosophical baggage—people may be asked to entertain more mundane scenarios designed to assess their beliefs. For example, they may be asked to consider a certain neurological disease, as in the case of frontotemporal dementia, and predict its psychological consequences (Strohminger & Nichols, 2015). They may be asked to assess how much the brain contributes to various mental constructs, including personality traits,

diego.fernandezduque@villanova.edu

cognitive processes, and even the sense of self (Fernandez-Duque & Schwartz, 2016; Johnson & Wellman, 1982). By breaking the mind/brain problem into smaller units, researchers aim to obtain a more exact assessment of people's lay beliefs.

Finally, another approach is to probe mind/brain beliefs implicitly. In other words, rather than asking people about the relation between mind and brain, one could ask people to perform an ostensibly unrelated judgment that offers a window into their lay beliefs. For example, one could ask participants to assess the quality of an explanation of a psychological phenomenon, as a way to test whether the presence of neuroscience information increases the perceived quality of the explanation (it does; Fernandez-Duque, Evans, Christian, & Hodges, 2015; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008). Alternatively, one could ask participants to assess the moral responsibility behind hypothetical moral transgressions, as a way to test whether providing neuroscience information increases deterministic explanations of behavior and reduces moral condemnation (it does; Monterosso, Royzman, & Schwartz, 2005). In sum, this implicit approach uses neuroscience information—that is, information about the brain—as the independent variable to explore its influence on judgments of mental life. To the extent that those influences are documented in research, they provide psychologists with evidence (albeit of the indirect variety) that a connection exists between brain and mind in people's lay theories. Obviously, there is a big explanatory gap between showing that a connection exists and providing a full description of such a relation. This limitation notwithstanding, the study of the neuroscience allure has sparked interest in its own right, due to its relevance for many disciplines, including mental health, education, and law.

## Overview of the Chapter

It is the last approach—the implicit assessment of beliefs about mind and brain—that constitutes the bulk of this chapter. Particular emphasis is given to the allure of neuroscience explanations. Next, I review the effect of neuroscience on judgments of responsibility and free will, as studied in the fields of psychology and law. This is followed by a review of folk beliefs on brain/mind/self, and a brief review of the developmental literature on this issue. I conclude the chapter by pointing out the importance of research on lay theories of mind/brain, arguing that different conceptualizations of the mind/brain relation have profound implications for public policy in mental health research and practice.

## The Allure of Neuroscience: A Brief History

Over the last quarter century, our understanding of the mind and the brain have undergone a revolution. At the center of those changes has been the development of new neuroimaging techniques that have allowed scientists to create maps

diego.fernandezduque@villanova.edu

connecting mental processes to their putative neural substrates. The mapping process is inferential and overwrought with statistical assumptions, but maps they are, pictures of brain activity where years ago there was only mind. Importantly, many of the new techniques are noninvasive, and therefore safe for use in humans. As a consequence, those uniquely human capacities that so long fascinated Descartes have, in the last two decades, become ripe for neuroscientific inquiry: the neural mechanisms of creativity, rational thought, morality, language, and the self are no longer out of bounds for scientific exploration.

The scientific advances of the last three decades have been accompanied by increased attention from the media and increased fascination by the public. The attention is deserved; cognitive neuroscience has provided great additional explanatory power to the mechanisms that underlie psychological processes. Sometimes, however, superfluous information is added that does not provide additional insight. What happens then? Are people fooled by cognitive neuroscience, or by the images we usually associate with it?

To start answering this question, McCabe and Castel (2008) had participants read a one-page summary of a cognitive neuroscience finding written for the popular press. This baseline condition was compared to experimental conditions in which the same neuroscience information was accompanied by either a functional MRI image or a bar chart. Participants rated the scientific reasoning most highly when the neuroscience explanation was paired with the fMRI. McCabe and Castel concluded that brain images conferred credibility to the neuroscience explanations. However, subsequent studies have failed to replicate these findings, and the current consensus in the field is that brain images have little to no effect on the perceived quality of neuroscientific explanations. However, it remains a possibility that neuroscience information—pictorial or text-based—might influence the perceived quality of psychological explanations. We turn to that literature next.

## Neuroscience Increases the Appeal of Psychological Explanations

What happens when dubious references to brain mechanisms are brought up to pseudo-explain a psychological phenomenon? In those instances, are audiences more accepting of neuro-gibberish than of regular gibberish, and if so, why? To start answering this question, Weisberg and collaborators asked participants to read vignettes about well-established psychological phenomena and their possible explanations (Weisberg et al., 2008). The description of the phenomena was always accurate, but the quality of the explanation was variable: sometimes the arguments were good while other times they were circular, a mere restatement of the phenomenon. Consider, for example, a vignette reporting that, in visual tasks, the patterns of response times were different for faces than for places. In that vignette, the good explanation said this happened “because people use different processes to

diego.fernandezduque@villanova.edu

recognize faces than they use to recognize places." In contrast, the circular explanation claimed it happened "because the participants' responses were contingent on whether they saw a face or a place on the screen". A second factor provided the critical manipulation: half of the vignettes included superfluous neuroscience sentences, while the other half did not. The neuroscience information was not wrong, it was simply irrelevant to the phenomenon it was trying to explain. In the above example, it said that "neuroscientists have shown that the extrastriate cortex—an area of the brain known to be involved in processing complex visual stimuli—is activated by pictures of faces and places." Despite its lack of relevance, such superfluous neuroscience sentences increased the perceived quality of circular explanations.

Weisberg's findings have been largely replicated by other labs, confirming that the 'allure of neuroscience' is conceptual rather than perceptual, meaning that neuroscience information is persuasive regardless of whether it is presented in form of brain images or neuroscientific text (Fernandez-Duque et al., 2015; Michael, Newman, Vuorre, Cumming, & Garry, 2013). This raises the question: why is neuroscience so alluring? One answer might be that neuroscience is a prestigious science that people trust. A different answer would be that neuroscience offers reductive explanations of psychological phenomena.

Unlike the social sciences, neuroscience is considered a 'real' science. When asked about the prestige of neuroscience, or about the gap between a neuroscience expert and a lay person, or about the scientific rigor of the discipline, undergraduate students always cluster neuroscience with other natural sciences and away from social sciences and psychology (Fernandez-Duque et al., 2015). Other studies show that people believe that biological explanations are more complex and more scientific than psychological explanations. This bias toward the natural sciences emerges as early as kindergarten, and vestiges of it can be observed in adulthood (Keil, Lockhart, & Schlegel, 2010). For example, the mere presence of a nonsense math equation increases the perceived quality of a scientific abstract (Eriksson, 2012), and the inclusion of a chemical formula increases the belief in a medication's efficacy (Tal & Wansink, 2014); just telling people that scientists understand a phenomenon is enough to increase people's judgment of their own understanding (Sloman & Rabb, 2016). In sum, there is little doubt that neuroscience is held in high regard as a science, and that scientific jargon often creates an illusion of understanding. Put these two facts together, and one might conclude that the allure of neuroscience is driven by its prestige.

However, if neuroscience's allure in explanations of psychological phenomena had to do with its status as a prestigious science, we would expect that gibberish from other "hard sciences" would also be alluring, provided that its relation to the psychological phenomena was not too far fetched; after all, such hard sciences—unlike psychology—are prestigious too. In our study, we explored this hypothesis by including irrelevant information from the hard sciences; for example, in the previously described vignette on visual processing of faces and places, participants read that "computational scientists have used spectrograms to show that pictures of faces and places convey a range of spatial frequencies." As in the case of

diego.fernandezduque@villanova.edu

neuroscience, the information was true but not particularly informative in explaining the psychological finding (Fernandez-Duque et al., 2015). The results of the study show convincingly that pseudo-explanations from other hard sciences are not as compelling as neuroscience pseudo-explanations (Fernandez-Duque et al., 2015).

What, then, explains the neuroscience allure? Although a definitive answer to this question has not yet been reached, we and others have speculated that the reason why neuroscience information is seen as more relevant than hard science or psychological counterparts is that in Western cultures, educated people conceptualize the brain as the engine of the mind (Fernandez-Duque & Schwartz, 2016; Hopkins, Weisberg, & Taylor, 2016). That is, participants in our studies may be conceptualizing the brain as the physical substrate that instantiates the mind, the structure to which psychological phenomena may one day be mapped or reduced.

The perspective we have been advocating so far in this section is that the allure of neuroscience for psychological phenomena stems from a lay theory according to which the brain is the engine of the mind, or put slightly differently, it is the next level of analysis below the mind. If this perspective is correct, then the allure of neuroscience would be just a special case of a more general heuristic, by which information is alluring if it sits at the level below the phenomenon of interest. According to this position, the most alluring information is reductive information. To test this hypothesis, Hopkins et al. (2016) extended the original paradigm to include other scientific disciplines, such as physics, chemistry, and biology. As predicted, they found that superfluous information was most alluring at the level just below the phenomena in need of explanation. In other words, people did prefer reductive explanations. As a caveat, it should be pointed out that the allure for the psychology/neuroscience pair was larger than for any other pair, leaving open the possibility of additional content specific influences above and beyond the allure of reductive explanations.

## Neuroscience Influences Judgments of Responsibility

In 1848, a railroad worker named Phineas Gage suffered a terrible accident when a metal rod exploded in his face, impaled him through cheek bone and skull, and in its way destroyed large parts of his frontal lobe. Gage survived, but his personality changed profoundly. Previously, he had been a conscientious worker, intelligent and well adapted. After the accident, he became “fitful, irreverent, indulging at times in the grossest profanities (which was not previously his custom), manifesting little deference for his fellows, impatient of restraint or advice when it conflicts with his desires” (Harlow, 1868). Friends and acquaintances would report that ‘Gage was no longer Gage’. Most of us would not blame Gage. Instead, we would assume that his behavioral outbursts and moral transgressions were outside of his control—that he did not have much choice—and conclude that he should not be held responsible, or at least not as responsible as someone with full mental capacity.

diego.fernandezduque@villanova.edu

Probably, we would not need to see a picture of his brain or his skull to reach these conclusions.

Our response to Gage's behavior nicely illustrates some of the folk beliefs regarding brain, moral responsibility, and free will. Over the last decade, researchers have begun to systematically assess such folk beliefs (Greene & Cohen, 2004). These studies are remarkably consistent in showing that neuroscience information does influence participants' judgments of moral responsibility. In other words, this literature on neuroscience and responsibility tells the very same story already presented regarding the allure of neuroscience and psychological phenomena.

In one of the first studies of its kind, Monterosso, Royzman, and Schwartz (2005) asked participants to read vignettes describing individuals who had committed a moral transgression. The authors varied whether the explanation for the transgression was neurobiological (e.g. "unusually high levels of a particular neurotransmitter") or experiential (e.g., "severely and brutally abused as a child"), reasoning that a neurobiological explanation would elicit a mechanistic view, whereas an experiential one would not. As predicted, the neurobiological explanation led to less blame than the experiential explanation, and to reduced ascription of willful control. In a follow-up study, the same two explanation types (neurobiological, experiential) were factorially crossed with the presence or absence of a neuroimage. Participants were asked to judge the extent to which the transgression was due to lack of moral character. By itself, the experiential justification led to larger moral condemnation than the neurobiological justification, but this difference disappeared when a neuroimage was attached. Participants who saw the brain image together with the experiential explanation probably took the brain to be the mechanistic mediator of the experiential account, and thus reduced the target's responsibility (Beall et al., 2013).

Similar findings have been obtained in experiments in the field of psychology and law. For example, when participants in a mock trial had to decide on a case of not guilty by reason of insanity, they had a tendency to find neuroscience-based evidence more persuasive than psychological evidence (Schweitzer & Saks, 2011). In another study, the presence of neuroscientific testimony reduced the likelihood of a death sentence verdict; the reduction occurred irrespective of whether the evidence consisted of a brain image or neuropsychological testimony alone (Greene & Cahill, 2012). In yet another study, participants in a mock trial had to decide on a case of not guilty by reason of insanity. Participants were biased toward a not guilty verdict by evidence of neuroscience lesion (in this case, an MRI image), as well as by evidence of abrupt onset (in this case, an episode of traumatic brain injury) (Gurley & Marcus, 2008). Both of these factors—neural evidence and abrupt onset—are also evident in the illustrative case of Phineas Gage. Finally, one study using US state judges as participants showed that expert testimony on the neurobiological mechanisms of psychopathy causes judges to consider those mechanisms as mitigating factors, thus leading to reduced criminal sentences (Aspinwall, Brown, & Tabery, 2012).

diego.fernandezduque@villanova.edu

In summary, reading or listening to evidence about the neural bases of human behavior leads people away from attributions of moral responsibility and away from retributive punishment. One account of these results is that learning about neuroscience highlights a mechanistic worldview in which free will is diminished and therefore actors should be held less blameworthy for their acts (Greene & Cohen, 2004; Shariff et al., 2014). That account rests in part on the assumption that ordinary folk deem neuroscientific explanations of behavior to be constraints on free will. Whether this is indeed the case is a matter of debate, which is discussed next.

## Neuroscience Influences Judgments of Free Will

In the philosophical literature, 'free will' is discussed by appeal to metaphysical concepts such as 'uncaused agency' (i.e., the ability of an agent to act without such act being caused by something else). Free will thus defined is challenged by a deterministic world, and this leads philosophers to all sorts of intellectual contortions to try to establish a coherent view (Roskies, 2006).[1] Research exploring the relation between neuroscience and free will philosophically defined has had limited success; there is great variability in people's judgments across experiments, and there are often internally inconsistent responses. For example, some studies suggest that people embrace both determinism *and* free will, a position known as *compatibilism* (Monroe & Malle, 2010; Nahmias, 2006; Nahmias, Shepard, & Reuter, 2014; Nichols & Knobe, 2007). According to this view, when morally evaluating an action, people state that even if the universe is fully deterministic, the actor could act differently. Other studies show instead that when presented with neuroscience claims that "free will does not exist because choices are caused by neural impulses" people reply by appealing to a different level of analysis, focusing on the agent to argue that "the person makes the neural impulses happen." This way, people seem to endorse the neuroscientific correlates of psychological states without committing to a deterministic view of them. Yet some other research suggests that determinism undermines free will in the abstract, but does not excuse wrongdoing in concrete cases (Nichols, 2011). In general, the sense one gets from reading this literature is that people lack stable notions of 'free will'.

As it turns out, 'free will' as understood by the common person is quite a different concept from the one developed by professional philosophers. When asked to define 'free will', ordinary folk do not refer to metaphysical criteria, rather, they provide a psychological account. People report that 'free will' consists of the ability to make choices consistent with one's desires, reasonably free of constraints; they sometimes also emphasize the reflective, deliberate aspect of it, that is, the

[1]These contortions include rejecting determinism to save free will (libertarians); conceding that the world is deterministic and thus acknowledging that that free will does not exist (hard determinists); or accepting determinism but still claim that free will is possible (i.e., compatibilism).

diego.fernandezduque@villanova.edu

forethought of weighting the pros and cons of the action (Monroe & Malle, 2010; Nahmias, 2016).

How does the folk concept of 'free will' relate to neuroscience and morality? The answer, at least hypothetically, is remarkably simple and powerful. To the extent that a neurological disorder disrupts one of the underlying psychological components of 'free will' (choice, desire, absence of external constraint, forethought), a reduction of 'free will' will ensue. In contrast, the existence of a normally functioning brain ought not pose a challenge to free will because a normal brain—by definition—has a correspondingly normal psychology. The threat of determinism that so much challenges the concept of 'free will' metaphysically defined simply vanishes once we adopt its common sense definition (Nahmias, 2006).

So far, I have described two competing conceptualizations of free will: one favored by philosophers which is metaphysically defined, and the other one favored by folk theory which is psychologically defined by 'choice'. What role do these two different conceptualizations of 'free will' play in judgments of moral responsibility? One way to answer this question is to pit the metaphysical notion of free will against the folk notion, in a $2 \times 2$ factorial design. In one such study, participants were divided into an experimental group that read a statement arguing against metaphysical free will ("all behavior is determined by brain activity, which in turn is determined by a combination of environmental and genetic factors") and a control group that read a neutral statement that made no reference to free will ("Oceans cover 71% of the earth's surface"). As expected, the experimental group reported less belief in metaphysical free will than the control group. Immediately after this manipulation, the participants watched a brief video in order to make a judgment of blame. In the 'choice' condition, the video depicted a person in a situation in which he could choose to steal money from a partner. In contrast, in the 'no choice' condition the amount of money taken was determined randomly. Participants who saw the 'choice' video assigned much more blame than participants in the 'no choice' condition. Importantly, the metaphysical manipulation had no effect (Monroe, Brady, & Malle, 2016). In another study, a vignette described a hypothetical study in which scientists were able to predict a person's future behavior based on her pattern of neural activation. Despite the scientists being able to predict in advance what the person would do (consistent with determinism), participants thought that the person was still exercising her free will (Nahmias et al., 2014). Only if the vignette described neuroscientists as bypassing the person's decision—that is, if it described the scientists as stimulating the patient's brain to actively manipulate her choice—did participants consider the patient to be deprived of free will.

In sum, the folk judgment of whether a person has 'free will' seems to depend on the mental states of the person: if her desires and choices are efficacious for causing action, then free will is affirmed, and moral responsibility assigned. This framework is useful in helping us reinterpret the literature on neuroscience and judgments of responsibility. We excuse Phineas Gage's moral transgressions not because our knowing of his brain lesion turns us into hard determinists skeptical of metaphysical

diego.fernandezduque@villanova.edu

free will, but rather because our knowing of his lesion reminds us of Gage's reduced mental capacities to make choices consistent with his desires and free of unreasonable constraints.

Alternatively, and on a more speculative note, when it comes to commitments regarding the relationships among mind, brain, and free will, people may be, in the words of philosopher Eddie Nahmias, 'theory-lite' (Nahmias, 2017); that is, people may have intuitions that are unstable and/or contradictory, without a reliable, strong commitment to dualism or materialism. After all, unlike professional philosophers, people can get through life without having to address metaphysical questions of the mind/brain relation. On the other hand, people cannot get through life without some form of moral theory to guide their judgments of blame and responsibility. Those judgments are dependent on free will in the psychological sense of the term. As neuroscience progresses, it seems likely that people will retain their belief-based model of free will and morality, and simply make ad hoc necessary adjustments to their 'theory-lite' metaphysics for the rare occasions in which such esoteric questions may arise.

## Interlude

Up to this point, the chapter has focused on the implicit assessment of beliefs about mind/brain relation. I have described the allure of neuroscience for psychological explanations, and the effect of neuroscience on judgments of responsibility and free will. The approach has been successful in showing that appeals to neuroscience do exercise an influence on judgments of mental activity. However, there is a big explanatory gap between showing that a connection exists and providing a full description of such a relation. One way to attempt closing that gap is to ask participants explicitly what they think about the relation between mind and brain. Doing so is not without challenges: people's responses to explicit questions are often more susceptible to wording artifacts, explicit questioning may trigger ad hoc answers that fail to align with the background beliefs regularly held, and answers may be altogether unreliable if people lack well-established theories. These limitations notwithstanding, listening to people's insights about their beliefs of mind/brain relation may enrich our understanding of such lay theories.

## Folk Beliefs About Brain and Mind

Developmental psychology has had a long-standing interest in understanding children's conceptual development. Thus, it comes as no surprise that much of the pioneering work on folk theory of mind and its relation to the brain can be traced back to developmental psychologists (e.g., Lillard, 1998; Johnson & Wellman, 1982). Such research has found that in Western cultures, the mind is often identified

diego.fernandezduque@villanova.edu

with the brain (Lillard, 1998). When asked "Do you need the brain to ____?", both adults and elementary school children endorse the view that the brain is necessary for all sorts of human psychological activities. These include emotions such as feeling sad or feeling curious, and senses such as hearing and seeing, but also cognitive acts like thinking and knowing, as well as reading and writing. It includes motor tasks like talking and walking and, in the case of adult participants, even involuntary tasks like coughing and blinking (Johnson & Wellman, 1982). In other words, when asked about the functions of mind and brain, elementary school children and adults alike treat the brain as responsible for the functions of the mind.

Both ninth graders and adults reject that the mind could exist in the absence of the brain, and both of them localize mind and brain in the head. However, ninth graders and adults say that, unlike the brain, the mind is nonmaterial and thus could not be seen nor touched even if the head were opened up. In contrast, young children seem to have a different ontology of the mind. For example, first graders conceive the mind as a material entity that could be seen and touched as much (or as little) as the brain; first graders also tend to construct mind and brain as independent entities, and claim that a mind could exist without a brain (Johnson & Wellman, 1982). Adults and older children also believe that the mind has temporal cohesiveness: they understand that it is the same mind which, encompassing various cognitive processes and states, is being used at different times (Johnson & Wellman, 1982).

## Folk Beliefs About Brain and Self

Interestingly, beliefs about spatial and temporal cohesiveness apply not only to the mind but also to the self. The cohesiveness of the self is nicely illustrated by Descartes' famous inference "I think, therefore I am". In this statement, Descartes assumes the existence of a self ("I") doing the thinking. More generally, both adults and children localize the *self* near the eyes (Anglin, 2014; Bertossa, Besa, Ferrari, & Ferri, 2008; Starmans & Bloom, 2012). Besides this conceptualization of the self as the experiencer (the "I"), there is also a conceptualization of the self as an object (the "me"). In that regard, adults think of themselves not as a disparate collection of thoughts and dispositions but rather as a cohesive unit somewhat stable over time, especially as it projects into the future (Moore, Lemmon, & Skene, 2001; Neisser, 1988; Quoidbach, Gilbert, & Wilson, 2013). In other words, the concept is temporally extended to also include the past self and the future self, with interesting asymmetries between the two. For example, adults of all ages think that their values, preferences, and personality traits, having evolved in the past, have now reached a stasis that protects them against further change (Quoidbach et al., 2013). For retrospective judgments, people favor downward comparisons, especially for the distant past. In one such study, college students were given a list of positive attributes, such as willingness to stand up for one's beliefs, or having good social skills. Students had to indicate the degree to which they possessed each attribute

diego.fernandezduque@villanova.edu

relative to their same-aged peers, on a scale from 0 (much less than most) to 10 (much more than most). Students assessed the self twice: first as they remembered it at age 16 and then as they knew it at present time; maybe not surprisingly, the rates of positive attributes at age 16 were substantially lower than at their current age. Replications at other ages ruled out an account based on poor adolescent skills (Wilson & Ross, 2001). Instead, the results are best explained by temporal self-appraisal theory, according to which people are motivated to enhance their perception of their current self. In pursuit of this goal, people implicitly make downward comparisons with their former self, as long as the former self is distant enough that it can be plausibly rejected from the current self-conception (Peetz & Wilson, 2008). Self-appraisal theory highlights that the concept of self is not immune to motivated cognition. This should not be a surprise, given that the concept of the self is in itself part of a lay theory informed by semantic knowledge about the mind, as well as by autobiographical memory (Neisser, 1988).

Besides the distinctions with the distant past, there are also distinctions between near and distant future in the conception of the self. For the near future, people adopt a concept of self that is mostly concrete, specific, and context dependent; but for the distant future, people favor instead an abstract self, closest to the *true* or essential self (Wakslak, Nussbaum, Liberman, & Trope, 2008). This latter distinction highlights an important point, namely that the concept of the self is hierarchically organized, with some traits being more central and others being more peripheral (Markus & Wurf, 1987; Sedikides, 1995).

The central self can be defined as the person you *truly* are (i.e., your *true* self) so that if you lacked those attributes you would be a different person; sometimes, the term 'core self' is used as synonymous. When asked to describe their 'central' self, people often make reference to moral traits, such as "honesty" and "kindness" (Fernandez-Duque & Schwartz, 2016; Goodwin, Piazza, & Rozin, 2014). In other words, when asked to describe who they *truly* are, what people volunteer are uniquely human traits, traits that Descartes believed did not belong in the brain. And while philosophers have long ago moved beyond Cartesian dualism, it remains a legitimate scientific contention that ordinary folk still hold to this belief (Bloom, 2004). According to this hypothesis, people would be willing to admit that the brain is the substrate of cognitive functions and many psychological traits, but would reserve a special nonmaterial place for traits that define who they truly are.

To test this hypothesis, we asked a group of people residing in the USA (recruited through Amazon's Mturk) to judge whether the brain was "more responsible for the CORE attributes of your self or for the PERIPHERAL attributes of your self" using a 100-point bipolar scale (Fernandez-Duque & Schwartz, 2016). We explained to the 172 participants that "the core self is who you truly are […] so that if you lacked those attributes you would be a different person" while the peripheral self included "things that describe you but don't define you […] so that if you didn't have those attributes, you would still be the same person." Contrary to the hypothesis, participants embraced the brain as the underlying substrate of their central self, that is, of who they *truly* are (Fernandez-Duque & Schwartz, 2016). We also asked another 210 participants about the neuroscience contribution toward 18

diego.fernandezduque@villanova.edu

different personality traits. Once again, the brain contribution for traits closer to the central self was deemed larger than for more peripheral traits. As expected, there was quite a bit of variability among traits, with perceived contributions of the brain ranging from 54% (for laziness) to 91% (for intelligence). In the future, studies probing a larger number of traits will help identify the trait attributes (e.g., volitional control, desirability, etc.) that best predict the perceived brain contribution.

This belief that the brain is the underlying substrate to people's *true* self and personality is also apparent in caregivers' reports of frontotemporal dementia patients. Brains affected by frontotemporal dementia are lesioned in areas similar to Phineas Gage's, and as a consequence these patients with frontotemporal dementia often exhibit similar behavioral and moral transgressions (Fernandez-Duque & Black, 2007; Fernandez-Duque, Hodges, Baird, & Black, 2010). Friends and families often report that the personal identity of the patient has changed since the start of the disease and that the patient "seems like a stranger" and "is not the same person underneath." At an intuitive level, these caregivers are endorsing the belief that changes to their loved ones' *true* self was brought about by pathological changes in their brains (Strohminger & Nichols, 2015). As such, it is an example of what in the folk psychiatry literature has been called neuro-essentialism, which is "the belief that brains and their abnormalities define and determine identity" (Haslam, 2011).

## The Possibility of Dualism

Many religious beliefs that are popular across the world depend on a dualist concept of mind/brain relation (or at the very least of the soul/brain relation). For example, beliefs about the afterlife require the existence of a nonmaterial substance separate from the body (Greely & Hout, 1999). Furthermore, there seems to be a clear positive correlation between popular dualism and other beliefs that seem dependent on it, such as beliefs in the afterlife, paranormal beliefs, and some religious beliefs (Fernandez-Duque & Schwartz, 2016; Riekki, Lindeman, & Lipsanen, 2013). Based on the evidence like this, as well as some of the developmental literature, some researchers have argued that children start as dualists, and become materialists only years later—if at all—through formal education (Bloom, 2004). According to this view, people learn in school, and through the internet and other media, that "the brain underlies the mind" the same way that people learn all sorts of strange, unintuitive scientific facts (Bloom, 2004). The evidence for and against dualism stems from various fronts and is described in detail in other chapters of this book (Haslam, Chapter "The Origins of Lay Theories: The Case of Essentialist Beliefs"; Forstmann & Burgmer, Chapter "Antecedents, Manifestations, and Consequences of Belief in Mind–Body Dualism").

However, in interpreting these data, and in interpreting data on folk theories of mind/brain more generally, it is important to keep in mind that any characterization of common sense beliefs about the mind/brain relation needs to contend with the fact that the concept of the 'mind' is not a monolithic construct but rather a multifaceted

diego.fernandezduque@villanova.edu

one; therefore it is possible—and even likely—that common sense beliefs about the mind may similarly include a constellation of different beliefs, with some psychological states deemed more brain based than others. For example, American adults tend to cluster mental states into two dimensions, an experiencing/feeling dimension that includes psychological states such as the feelings of hunger, fear, and joy and an agency/cognitive dimension that includes psychological states such as self-control, morality and memory (Gray, Gray, & Wegner, 2007). A biological brain appears necessary for experiencing things, such as hunger, joy, or pleasure, as people do not attribute those experiences to God or a robot. In contrast, for agency, a biological brain seems neither necessary nor sufficient: God and robots are deemed high on agency while newborns and frogs are denied it (Gray, Gray, & Wegner, 2007). Similarly, people resist attributing experiences to brainless corporations ("Acme is feeling pain") but accept attributions of agency to them ("Acme Corp plans to change its corporate image" Knobe & Prinz, 2008).

## Why Is It Imperative to Understand Lay Theories of Mind/Brain?

The issues discussed in this chapter are important to understand not only because they enrich our description of how humans categorize and conceptualize the world —in the tradition of past research on folk physics (McCloskey & Kohl, 1983), folk biology (Carey, 1985), and folk mentalizing (Wellman, Cross, & Watson, 2001)— but maybe more importantly because the decisions humans make, and the social worlds that they construct, derive directly from the lay theories and beliefs they hold regarding those worlds. Thus, different conceptualizations of the mind/brain relation should have profound implications for public policy in mental health research and practice. By better understanding those conceptualizations we might be able to modify them, and in doing so, we may be able to modify our destiny. These aspirations sound lofty and vague, so some concrete elaboration is in order.

Let us start by stating the obvious: the natural world does not care about the theories humans create to explain it. Alchemy in the seventeenth century may have proposed a theory to turn copper into gold, but no amount of theorizing was ever going to make that happen. By contrast, the social world is quite susceptible to the theorizing we humans do in order to explain it. The distinguished psychologist Barry Schwartz has illustrated this point in his explanation of why we work (Schwartz, 2015). According to Adam Smith, the father of free market economic theory, we work for pay, nothing more and nothing less. A workplace in which workers' only motivation is thought to be monetary is likely to be designed lacking any other sources of motivation; after all, why waste resources promoting workers' sense of accomplishment, or creating a supportive social life in the workplace, if we *know* that workers only care about their paycheck? In such a devoid environment, workers would not find any reason to work other than their salary, and when asked, they will confirm our initial theory.

diego.fernandezduque@villanova.edu

If this analysis is correct, then we *discover* the natural world but we *create* the social world. When it comes to lay theories of the mind/brain relationship, that social world includes, among other things, the treatment of mental disorders, the funding priorities for mind/brain research, and the implementation of our legal and educational systems.

Consider, for example, the treatment of mental disorders. When mental disorders, such as ADHD or generalized anxiety disorder are explained by appeal to biological information, people become overly pessimistic about their prognosis (Lebowitz & Ahn, 2014; Lebowitz, Pyun, & Ahn, 2014). When therapists hear of mental diseases such as depression and OCD in biological terms, they become less sympathetic toward the patients (Lebowitz & Ahn, 2014). Therapists' beliefs are quite malleable, so that those with a medical degree are more inclined to think of the disorders as medically based (Kim, Ahn, Johnson, & Knobe, 2016). More arbitrary biases are present too; when the disorder is described abstractly in terms of symptoms, therapists think of it as biologically based, but when described concretely in relation to an individual patient, the same therapists become more inclined to think of the disorder as psychologically based, and less susceptible to medical treatment; this is true even for those therapists who are medically trained (Lebowitz, Rosenthal, & Ahn, 2016). Therapists and laypersons conceptualize mental disorders along a single continuum that spans from disorders considered highly biological (e.g., autism) to disorders considered highly psychosocial (e.g., adjustment disorders), thus ignoring the quite likely scenario of dual contributions from biological and social factors (Ahn, Proctor, & Flanagan, 2009). The goal of these examples is not to dwell on this very interesting literature (for a deep analysis, see Furnham, chapter "How Lay Theories Influence our Mental Health") but rather to illustrate the claim that our folk theories of the mind/brain relation have a profound impact on how we approach and try to solve the problems in front of us, both in terms of clinical practice and of public policy.

Another illustrative example comes from the funding priorities for mind/brain research at the United States' National Institute of Mental Health, which in the last decade, under the directorship of Tom Insel, a neuroscientist known for his work on hormonal control of monogamy in mammals, has redirected its focus away from social science and toward neuroscience, where it is now almost exclusively focused (Markovitz, 2016). It seems reasonable to speculate that the folk theory at the helm regarding the relation between mind and brain has been at least partly responsible for those changes. It also seems reasonable to ask whether those changes in priorities reflect the values of the citizenry, as expressed by folk beliefs. For example, should the taxonomy of mental disorders be organized by identifying symptom clusters, or should it instead be built bottom up from genes and neurobiology? Should promising behavioral therapies receive funding for effectiveness testing, or lose such funding due to the lack of a neurological correlate in their proposed mechanism? Should the effectiveness of potential treatments for Alzheimer's disease be judged based on their ability to remove the biological substrate of the disease (i.e., presumably plaques) or by their ability to improve behavior (e.g., episodic memory)? As the preceding discussion makes clear, there is a lack of

diego.fernandezduque@villanova.edu

consensus among scientists about the level of analysis at which mental disorders ought to be conceptualized and treatment delivered; the answer to such questions depends to a great extent on the particular disease under discussion, and often the most effective treatments combine interventions at both neural and psychological levels (i.e., drugs and talk therapy).

In the field of education, the situation is quite different. Although there has been a fair amount of hype surrounding neuroscience and education for the last 20 years, the level of analysis at which educational gains are maximized is undoubtedly psychological rather than neurological (Bowers, 2016). This should not be surprising; after all, the primary outcomes of education are behavioral: we want children to learn to read, do math, develop critical thinking skills, and so forth. And unlike mental disorders, the treatment options are exclusively behavioral: in order to foster children's phonemic decoding, we sound out letters, in order to foster a number sense, we draw a number line, and so forth. At most, the potential of neuroscience to affect education is likely limited to low-level behaviors, such as reading, rather than more complex behaviors such as collaborating with a classmate or writing an essay. Nonetheless, how likely are scientific theories of mind/brain relation to influence education in years to come? To start answering this question, it is helpful to start with a brief history of the mind/brain relation in science.

In the early 1800s, phrenologists had aimed to divide the mind into its constituent mental faculties but had failed spectacularly, due to a lack of empirical rigor and a penchant for ill-conceived categorization. By the late 1800s, Paul Broca had overcome some of these limitations by discovering that speech could be mapped to a specific region of the brain (Broca, 1861; Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007). For many decades afterwards, the prevalent paradigm remained trying to relate large complex task capacities—speech, memory, motor control—to similarly large brain regions (Thiebaut de Schotten et al., 2015).

But in the '70s, cognitive psychologists started to break down those large cognitive capacities into smaller mental operations. To achieve this, they administered relatively simple tasks, contrasted nearly identical experimental conditions, and measured response times with millisecond precision. This way, cognitive psychologists were able to isolate what they referred to as "elementary mental operations," the building blocks from which complex cognitive tasks are made (Posner, 1978). In the following decades, proponents of these "chronometric explorations of the mind" would convincingly argue that such elementary mental operations constituted the appropriate level of analysis at which to map mind and brain (Posner & Raichle, 1994). The idea became the main tenet of the new field of cognitive neuroscience, and helped cognitive neuroscience move past not only phrenology's but also Paul Broca's conception of the mind.

This scientific conceptualization of the mind/brain relation has provided cognitive neuroscientists with the necessary theoretical models to uncover the neural bases of reading (Dehaene, 2009), mental calculation (Dehaene, 2011), and working memory, as well as many other cognitive processes, each with its own set of elemental mental operations. As a consequence, sophisticated neural models of dyslexia and dyscalculia have been developed over the last 20 years,

diego.fernandezduque@villanova.edu

and neuroimaging studies can now predict above chance which children will develop dyslexia, and which of them will benefit from treatment, thus opening the door to personalized educational treatment (Butterworth, Varma, & Laurillard, 2011; Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015). Equally important, the models have had an impact on educational policy: based on the neuroscience evidence, the National Institute of Child Health and Human Development now defines dyslexia as "a specific learning disability that is neurobiological in origin." As this example nicely illustrates, the scientific theory of the mind/brain relation has already started to influence some aspects of educational policy and practice, although the extent of such influence in the future remains to be seen.

In this section, I have argued that a better understanding of current theories of the mind/brain relation is important, not only as a basic scientific endeavor, but also as a tool for public policy and practice. The impact of such theories is found in areas as diverse as mental health, education, and science funding, as well as in the field of psychology and law. Unresolved issues remain, such as the potential gap between the folk theory and the expert theory, and how to adjudicate in cases in which folk theory and expert theory disagree.

## Summary

In this chapter, I have tried to answer a deceptively simple question: Why is neuroscience so alluring? But to ask this question is to ask about current lay theories of the mind/brain relation. I started by reviewing the allure of neuroscience explanations for psychological phenomena. I showed that although neuroscience is a prestigious science, this does not explain its allure. Nor can the allure of neuroscience explanations be explained by the neuroimages that often accompany them. Instead, the allure stems from a folk theory of the mind that conceptualizes the brain as the engine of the mind. Neuroscience is alluring for explaining psychological phenomena because of its reductive appeal.

Next, I reviewed the effect of neuroscience on judgments of responsibility and free will. This is another way to ask about the folk theory of the mind/brain relation: how the brain contributes to uniquely human attributes such as free will and morality. I showed that although ordinary folk shy away from free will and moral condemnation when actions are couched in terms of brain function, the reason for this is not that people think of free will as something that cannot exist in a fully deterministic material world. Rather, people conceptualize free will as the ability to make choices consistent with one's desires, reasonably free of constraints. To the extent that a neurological disorder disrupts these underlying psychological components of 'free will', people attribute a reduction of 'free will'. Thus, we excuse the moral transgressions of patients who, like Phineas Gage, have abrupt brain lesions not because our knowing of his brain lesion turns us into hard determinists skeptical of metaphysical free will, but rather because our knowing of his lesion reminds us of Gage's reduced mental capacities to make choices consistent with his

diego.fernandezduque@villanova.edu

desires and free of unreasonable constraints. In contrast, the existence of a normally functioning brain ought not pose a challenge to free will because a normal brain—by definition—has a correspondingly normal psychology.

The next stop on our tour of uniquely human capacities and their relation to the brain was the notion of the self. Do ordinary folk believe that their brains are who they *truly* are? Recent evidence suggests that they do, at least for Western educated cultures. The chapter also briefly discussed mind–body dualism and neuro-essentialism, topics that get the attention they deserve in other chapters of the book (Haslam, Chapter "The Origins of Lay Theories: The Case of Essentialist Beliefs"; Forstmann & Burgmer, Chapter "Antecedents, Manifestations, and Consequences of Belief in Mind–Body Dualism").

Finally, I concluded by exploring possible policy implications of the allure of neuroscience and current theories of the brain/mind relation. I have argued that the impact of such theories is found in areas as diverse as mental health, education, and science funding. As such, I hope they help illustrate the profound implications that differences in the conceptualizations of the mind/brain relation may have, not only for our understanding of human cognition but perhaps more importantly, for human society.

## References

Ahn, W. K., Proctor, C. C., & Flanagan, E. H. (2009). Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. *Cognitive Science, 33* (2), 147–182. doi:10.1111/j.1551-6709.2009.01008.x

Anglin, S. M. (2014). I think, therefore I am? Examining conceptions of the self, soul, and mind. *Consciousness and Cognition, 29,* 105–116. doi:10.1016/j.concog.2014.08.014

Aspinwall, L. G., Brown, T. R., & Tabery, J. (2012). The double-edged sword: does biomechanism increase or decrease judges' sentencing of psychopaths? *Science, 337*(6096), 846–849. doi:10.1126/science.1219569

Beall, E., Meindl, P., Iyer, R., Monterosso, J., Graham, J., & Schwarz, B. (2013). *Biology or bad character? Brain scan images increase determinism and decrease moral attributions for criminal acts*. Paper presented at the Annual Conference of the Society for Personality and Social Psychology, New Orleans, LA.

Bering, J. M., & Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology, 40*(2), 217–233. doi:10.1037/0012-1649.40.2.217

Bertossa, F., Besa, M., Ferrari, R., & Ferri, F. (2008). Point zero: A phenomenological inquiry into the seat of consciousness. *Perceptual and Motor Skills, 107*(2), 323–335.

Bloom, Paul. (2004). *Descartes' baby: How the science of child development explains what makes us human*. New York: Basic Books.

Bowers, J. S. (2016). The practical and principled problems with educational neuroscience. *Psychological Review, 123*(5), 600–612. doi:10.1037/rev0000025

Broca, P. (1861). Remarks on the seat of the faculty of articulated language, following an observation of aphemia (loss of speech). Remarques sur le siége de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin de la Société Anatomique, 6,* 330–357.

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science, 332*(6033), 1049–1053. doi:10.1126/science.1201536

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, Mass.: MIT Press.

diego.fernandezduque@villanova.edu

Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. New York: Viking.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed.). New York: Oxford University Press.

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 Science, 1157,* 1–9. doi:10.1111/j.1749-6632.2008.04117.x

Descartes, R. (1641) (1984/1641). Meditations on first philosophy. In (Cottingham, Trans.) R. J. Stoothoff & D. Murdoch (Eds.), *The philosophical writings of rené descartes* (Vol. 2, pp. 1–62). Cambridge: Cambridge University Press.

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. *Brain, 130*(Pt 5), 1432–1441. doi:10.1093/brain/awm042

Eriksson, K. (2012). The nonsense math effect. *Judgment and Decision Making, 7*(6), 746–749.

Fernandez-Duque, D., & Black, S. E. (2007). Metacognitive judgment and denial of deficit: Evidence from frontotemporal dementia. *Judgment and Decision Making, 2*(6), 359–370.

Fernandez-Duque, D., Evans, J., Christian, C., & Hodges, S. D. (2015). Superfluous neuroscience information makes explanations of psychological phenomena more appealing. *Journal of Cognitive Neuroscience, 27*(5), 926–944. doi:10.1162/jocn_a_00750

Fernandez-Duque, D., Hodges, S. D., Baird, J. A., & Black, S. E. (2010). Empathy in frontotemporal dementia and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 32*(3), 289–298. doi:10.1080/13803390903002191

Fernandez-Duque, D., & Schwartz, B. (2016). Common sense beliefs about the central self, moral character, and the brain. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.02007

Gabrieli, J. D., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron, 85*(1), 11–26. doi:10.1016/j.neuron.2014.10.047

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology, 106*(1), 148–168. doi:10.1037/a0034726

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315* (5812), 619. doi:10.1126/science.1134475

Greely, A. M., & Hout, M. (1999). Americans' increasing belief in life after death: Religious competition and acculturation. *American Sociological Review, 64,* 813–835.

Greene, E., & Cahill, B. S. (2012). Effects of neuroimaging evidence on mock juror decision making. *Behaviral Sciences and the Law, 30*(3), 280–296. doi:10.1002/bsl.1993

Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 359* (1451), 1775–1785. doi:10.1098/rstb.2004.1546

Gurley, J. R., & Marcus, D. K. (2008). The effects of neuroimaging and brain injury on insanity defenses. *Behavioral Sciences and the Law, 26*(1), 85–97. doi:10.1002/bsl.797

Harlow, J. (1868). Recovery from the passage of an iron bar through the head. *Massachusetts Medical Society, 2,* 327–347.

Haslam, N. (2011). Genetic essentialism, neuroessentialism, and stigma: commentary on Dar-Nimrod and Heine (2011). *Psychological Bulletin, 137*(5), 819–824. doi:10.1037/a0022386

Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition, 155,* 67–76. doi:10.1016/j.cognition.2016.06.011

Johnson, C. N., & Wellman, H. M. (1982). Children's developing conceptions of the mind and brain. *Child Development, 53*(1), 222–234.

Keil, F. C., Lockhart, K. L., & Schlegel, E. (2010). A bump on a bump? Emerging intuitions concerning the relative difficulty of the sciences. *Journal of Experimental Psychology: General, 139*(1), 1–15. doi:10.1037/a0018319

diego.fernandezduque@villanova.edu

Kim, N. S., Ahn, W. K., Johnson, S. G., & Knobe, J. (2016). The influence of framing on clinicians' judgments of the biological basis of behaviors. *Journal of Experimental Psychology: Applied, 22*(1), 39–47. doi:10.1037/xap0000070
Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences, 7,* 67–83.
Lebowitz, M. S., & Ahn, W. K. (2014). Effects of biological explanations for mental disorders on clinicians' empathy. *Proceeding National Academy of Science U S A, 111*(50), 17786–17790. doi:10.1073/pnas.1414058111
Lebowitz, M. S., Pyun, J. J., & Ahn, W. K. (2014). Biological explanations of generalized anxiety disorder: Effects on beliefs about prognosis and responsibility. *Psychiatry Services, 65*(4), 498–503. doi:10.1176/appi.ps.201300011
Lebowitz, M. S., Rosenthal, J. E., & Ahn, W. K. (2016). Effects of biological versus psychosocial explanations on stigmatization of children with ADHD. *Journal of Attentional Disorders, 20* (3), 240–250. doi:10.1177/1087054712469255
Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin, 123*(1), 3–32.
Markovitz, J. C. (2016). There is such a thing as too much neuroscience. *New York Times*, October 14. Retrieved from http://www.nytimes.com/2016/10/15/opinion/theres-such-a-thing-as-too-much-neuroscience.html
Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38,* 299–337.
McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition, 107*(1), 343–352. doi:10.1016/j.cognition.2007.07.017.
McCloskey, M., & Kohl, D. (1983). Naive physics: The curvilinear impetus principle and its role in interactions with moving objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(1), 146–156.
Michael, R. B., Newman, E. J., Vuorre, M., Cumming, G., & Garry, M. (2013). On the (non) persuasive power of a brain image. *Psychonomic Bulletin and Review, 20*(4), 720–725. doi:10.3758/s13423-013-0391-6
Monroe, A. E., Brady, G., & Malle, B. F. (2016). This isn't the free will worth looking for: General free will beliefs do not influence moral judgments, agent-specific choice ascriptions do. *Social Psychological and Personality Science, 27,* 100–108. doi:10.1177/1948550616667616
Monroe, A. E., & Malle, B. F. (2010). From uncaused will to conscious choice: The need to study, not speculate about, people's folk concept of free will. *Review of Philosophy and Psychology, 1,* 211–224.
Monterosso, J., Royzman, E. B., & Schwartz, B. (2005). Explaining away responsibility: Effects of scientific explanation on perceived culpability. *Ethics and Behavior, 15*(2), 139–158. doi:10.1207/s15327019eb1502_4
Moore, C., Lemmon, K., & Skene, K. (Eds.). (2001). *The self in time: Developmental Perspectives*. Mawah, NJ: Erlbaum.
Nahmias, E. (2006). Folk fears about freedom and responsibility: Determinism vs. reductionism. *Journal of Cognition and Culture, 6*(1–2), 215–237.
Nahmias, E. (2016). Free will as a psychological accomplishment. In D. Schmidtz & C. Pavel (Eds.), *The Oxford handbook of freedom*: OUP.
Nahmias, E. (2017). Your brain as the source of free will worth wanting: Understanding free will in the age of neuroscience. In G. Caruso & O. Flanagan (Eds.), *Neuroexistentialism: Meaning, morals, and purpose in the age of neuroscience*. Oxford: Oxford University Press.
Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology, 18*, 561–584.
Nahmias, E., Shepard, J., & Reuter, S. (2014). It's OK if 'my brain made me do it': People's intuitions about free will and neuroscientific prediction. *Cognition, 133*(2), 502–516. doi:10.1016/j.cognition.2014.07.009

diego.fernandezduque@villanova.edu

Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology, 1*(1), 35–59.
Nichols, S. (2011). Experimental philosophy and the problem of free will. *Science, 331*(6023), 1401–1403. doi:10.1126/science.1192931
Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs, 41*(4), 663–685.
Peetz, J., & Wilson, A. E. (2008). The temporally extended self: The relation of past and future selves to current identity, motivation, and goal pursuit. *Social and Personality Psychology Compass, 2,* 2090–2106.
Posner, M. I. (1978). *Chronometric explorations of mind: The third Paul M. Fitts lectures, delivered at the University of Michigan, September 1976*. Hillsdale, N.J.: L. Erlbaum Associates
Posner, M. I., & Raichle, M. E. (1994). *Images of mind*. New York: Scientific American Library.
Quoidbach, J., Gilbert, D. T., & Wilson, T. D. (2013). The end of history illusion. *Science, 339* (6115), 96–98. doi:10.1126/science.1229294
Racine, E., Bar-Ilan, O., & Illes, J. (2005). fMRI in the public eye. *Nature Review Neuroscience, 6* (2), 159–164. doi:10.1038/nrn1609
Riekki, T., Lindeman, M., & Lipsanen, J. (2013). Conceptions about the mind-body problem and their relations to afterlife beliefs, paranormal beliefs, religiosity, and ontological confusions. *Advances in Cognitive Psychology, 9*(3), 112–120. doi:10.2478/v10053-008-0138-5
Roskies, A. (2006). Neuroscientific challenges to free will and responsibility. *Trends Cogn Sci, 10* (9), 419–423. doi:10.1016/j.tics.2006.07.011
Schwartz, B. (2015). *Why we work* (First TED books hardcover edition September 2015. ed.).
Schweitzer, N. J., & Saks, M. J. (2011). Neuroimage evidence and the insanity defense. *Behavioral Sciences and the Law, 29*(4), 592–607. doi:10.1002/bsl.995
Sedikides, C. (1995). Central and peripheral self-conceptions are differentially influenced by mood: tests of the differential sensitivity hypothesis. *Journal of Personality and Social Psychology, 69*(4), 759–777.
Shariff, A. F., Greene, J. D., Karremans, J. C., Luguri, J., Clark, C. J., Schooler, J. W., et al. (2014). Free will and punishment: A mechanistic view of human nature reduces retribution. *Psychological Science, 25*(8), 1563–1570. doi:10.1177/0956797614534693
Sloman, S. A., & Rabb, N. (2016). Your understanding is my understanding: Evidence for a community of knowledge. *Psychological Science, 27*(11), 1451–1460. doi:10.1177/0956797616662271
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*(1), 5–23.
Starmans, C., & Bloom, P. (2012). Windows to the soul: Children and adults see the eyes as the location of the self. *Cognition, 123*(2), 313–318. doi:10.1016/j.cognition.2012.02.002
Strohminger, N., & Nichols, S. (2015). Neurodegeneration and Identity. *Psychological Science, 26* (9), 1469–1479. doi:10.1177/0956797615592381
Tal, A., & Wansink, B. (2014). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*. doi:10.1177/0963662514549688
Thiebaut de Schotten, M., Dell'Acqua, F., Ratiu, P., Leslie, A., Howells, H., Cabanis, E., … Catani, M. (2015). From Phineas Gage and Monsieur Leborgne to H.M.: Revisiting disconnection syndromes. *Cerebral Cortex, 25*(12), 4812–4827. doi:10.1093/cercor/bhv173
Wakslak, C. J., Nussbaum, S., Liberman, N., & Trope, Y. (2008). Representations of the self in the near and distant future. *Journal of Personality and Social Psychology, 95*(4), 757–773. doi:10.1037/a0012939
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. doi:10.1162/jocn.2008.20040
Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development, 72*(3), 655–684.
Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology, 80*(4), 572–584.

diego.fernandezduque@villanova.edu