

Social Neuroscience and the Soul's Last Stand

Joshua D. Greene

Version 11/06

To appear in Social Neuroscience: Toward Understanding the Underpinnings of the Social Mind

A. Todorov, S. Fiske, and D. Prentice, Eds. Oxford University Press.

“Eventually, as brain imaging is refined, the picture may become as clear and complete as those see-through exhibitions, at auto shows, of the inner workings of the internal combustion engine. At that point it may become obvious to everyone that all we are looking at is a piece of machinery, an analog chemical computer, that processes information from the environment. ‘All,’ since you can look and look and you will not find any ghostly self inside, or any mind, or any soul.”

--Tom Wolfe, “Sorry, But Your Soul Just Died” *Forbes*, 1996

Most people are dualists (Bloom, 2004). Intuitively, we think of ourselves not as physical devices, but as immaterial minds or souls housed in physical bodies.

Most experimental psychologists and neuroscientists disagree, at least officially.

The modern science of mind proceeds on the assumption that the mind is simply

what the brain does. We don't talk much about this, however. We scientists take the mind's physical basis for granted. And among the general public, it's a touchy subject. So why bring it up?

We scientists, of course, have our own touchy subjects. One of them concerns the value of neuroscientific research in psychology. Many argue that neuroscience, and brain imaging in particular, is highly overrated (Uttal, 2003). Paul Bloom (2006) attributes the seductive power of neuroscience to our intuitive dualism: "We intuitively think of ourselves as non-physical, and so it is a shock, and endlessly interesting, to see our brains at work in the act of thinking." Bloom, like many experimental psychologists, worries that we are spending millions of dollars on flashy experiments that do little to expand our knowledge of the mind, but that instead prompt us to contemplate our ontological navels.

I believe that our ontological navels are in desperate need of contemplation. On some level we appreciate the urgency of this enterprise, which is why so many psychologists—psychologists who are ardent non-dualists—are fascinated by neuroscience. But the dominant conception of psychological research and its aims obliges us to regard our fascination as an irrelevant distraction. We are forced into a kind of doublethink by which our deeper motivations are at odds with our official reasons for doing what we're doing. Like all scientists, neuroscientists who study mental phenomena are uncovering details. And as scientists we are supposed to be able to say why it matters if the details turn out one way rather than another. But when we try to explain why neuroscientific details matter to psychology, our work often sounds

like either an overpriced substitute for more traditional behavioral research or a plodding exercise in “brain mapping.” What, then, are we really trying to do?

And is it worth doing?

What we really want, I think, is to see the mind’s clockwork, “as clear and complete as those see-through exhibitions at auto shows.” That’s not all we’re after, of course. We’d like to cure diseases and do other patently useful things. But the promise of useful applications is not what fascinates us. Our fascination is existential. We are hooked on the idea of understanding ourselves in transparently mechanical terms. But a strange feature of this impulse to see the mind’s clockwork is that, so far as this impulse is concerned, the clockwork’s details are almost irrelevant. We don’t care how it works, exactly. We just want to see it in action. Is that foolish? I don’t think so. On the contrary, when we think about how our minds work more generally, this bare yearning to perceive the mechanical details of our minds, whatever they happen to be, makes perfect sense.

There are different ways of knowing (Kahneman, 2003; Lieberman, Gaunt, Gilbert, & Trope, 2002). It’s one thing to know something intellectually, to believe it in a thin, abstract sort of way—to say “yes” when asked if it’s true and not be lying. It’s quite another thing to know something in a deep way, to have one’s knowledge woven into the fabric of one’s worldview, guiding one’s thoughts and actions implicitly. Take, for example, the events of *9-11*. Most Americans were shocked that such a thing could happen. But why were people so surprised? Eight years earlier Islamic terrorists attempted to destroy the very same buildings

and nearly succeeded. Between 1993 and 2001, every rational person “knew” that America was vulnerable to a large-scale terrorist attack. But it took *9-11* to make people really *know*. To take another gruesome example, consider the need people sometimes feel to see the body of someone who has died. After reading the telegram, you may “know” that your long-lost brother is dead, but there is a kind of closure that comes only with seeing the body. The opposite happens at the movies. When the star-crossed lovers, locked in each others’ arms, tumble tragically into the lava pit, you “know” that it’s only a movie, but tell that to your thumping heart, your tearing eyes, and the parts of your brain that control them. All of this makes perfect evolutionary sense. The basic structure of our brains was in place long before we acquired the ability to use language, and with it the ability to acquire beliefs independent of sensory experience. There is a reason why we humans, who specialize in believing in things unseen, still insist that “seeing is believing.”

Our self-knowledge may be similarly fractured. Officially, we scientists already know (or think we know) that dualism is false and that we are simply complex biological machines. But insofar as we know this, we know this in a thin, intellectual way. We haven’t *seen* the absence of the soul. Rather, we have inferred its absence, based on the available evidence and our background assumptions about what makes one scientific theory better than another. But to truly, deeply believe that we are machines, we must see the clockwork in action.

We've all heard that the soul is dead. Now we want to see the body. This is what modern neuroscience promises to deliver, and it is no small thing.

One may argue that achieving a deeper understanding of ourselves is important in itself, on a par with understanding how the universe began and how life first arose on Earth. But I wish to make a more practical argument for deeper self-knowledge. Like a handful of others (Bloom, 2005; Dawkins, 2006; Dennett, 2006; Harris, 2004), I believe that our intuitive dualism causes a lot of problems. And if anything can talk us out of our dualist tendencies it is neuroscience—more specifically *social neuroscience*. According to Wegner and Gilbert (2000), social psychology has evolved from being a fairly circumscribed science of human social interaction into a sprawling science encompassing all of human subjective experience. If that is right, then the mission of social neuroscience, as the offspring of social psychology and neuroscience, is to understand all of human subjective experience in physical terms. The rise of social neuroscience is the demise of the soul.

My aim in this article is to consider the broader implications of social neuroscience, so conceived. While I am an unabashed enthusiast, I agree with critics who say that there is a real danger of our wasting precious time and money on misguided research. The challenge, as I see it, is to achieve the kind of short-term incremental progress that journals and funding agencies demand, while at the same time honoring our broader, and all but unspoken, philosophical mission. The key, I think, is to make our task the *functional decomposition of the brain*, i.e. breaking down complex psychological processes into simpler

processes that are associated with different parts of the brain. This is not a new idea. It's not even close to being a new idea. It's what the best cognitive neuroscientists and, more recently, social neuroscientists have been doing all along. But it's an idea that many people don't seem to get, especially people who come to social neuroscience without training in experimental psychology. (Witness the proliferation of scientifically under-motivated brain imaging experiments (Cacioppo et al., 2003).) And because social neuroscience is so inherently fascinating to our dualist minds, it's possible to get away with not getting it. In what follows I will argue that functionally decomposing the social brain is a worthy thing to do in the short-term, and perhaps one of the most worthy things that we social scientists could ever do in the long-term. As I make my case, I will use my own work on moral judgment as an illustrative example. There may be bad reasons for doing this (laziness, egocentricism), but there is at least one good reason. When it comes to undoing dualism, the neuroscientific study of moral judgment occupies a unique position.

These days, even the most ardent dualists recognize that we have brains and that our brains must do *something*. In recent decades we've learned that brains do many things that are historically within the province of the soul: perception, memory, the production and comprehension of language, etc. The soul has, as it were, "outsourced" these operations to the brain. This outsourcing process, still ongoing, raises a question: How many of the soul's functions can be taken up by the brain before the soul is completely out of a job? In other words, what is the soul's "core competence?" The answer, I believe, is *moral judgment*.

After all, in many religious traditions it is the quality of a soul's moral judgment and character that determines where it ends up, either permanently or on the next go-round. Thus, if the soul is not in the moral judgment business, it's not in any business at all. And, thus, what it would take to send the soul packing for good is a purely physical account of how the human mind does its moral business. If our goal is to determine once and for all whether there's a soul in there, there's no better place to start than with the neuroscience of moral judgment.

The Dual-Process Model of Moral Judgment

Consider the following moral dilemma, which we'll call the *crying baby* case:

It's wartime, and you and some of your fellow villagers are hiding from enemy soldiers in a basement. Your baby starts to cry, and you cover your baby's mouth to block the sound. If you remove your hand, your baby will cry loudly, the soldiers will hear, and they will find you and the others and kill everyone they find, including you and your baby. If you do not remove your hand, your baby will smother to death. Is it okay to smother your baby to death in order to save yourself and the other villagers?

This is a difficult question. People take a relatively long time to answer, and there is no consensus about what the right answer is (J. D. Greene, Nystrom, Engell, Darley, & Cohen, 2004). Why is this question so difficult? And what's going on in people's heads when they're deciding?

According to my dual-process theory (Greene et al. 2001, 2004), it goes something like this: On the one hand, we have an intuitive emotional response to the thought of smothering one's own baby (or anyone else) that makes us say, "No! It's wrong!" On the other hand, there's a different voice in our heads, a more dispassionate and controlled voice that says, "But there's nothing to be gained and much to be lost by not acting. The baby will die no matter what. You ought to save yourself and the others." These two voices, the intuitive emotional voice and the controlled "cognitive"¹ voice, fight it out in your head, until one of them wins and you render your judgment. This theory, despite its introspective plausibility, is at odds with two leading schools of thought in moral psychology, one of which denies that emotions play an important role in the moral judgments

¹ In some cases the word "cognitive" refers to information processing in general (as in "cognitive science"), while in other cases it refers to a kind of information processing that is contrasted with emotional or affective processes. Here and elsewhere I place "cognitive" in quotation marks to indicate the latter usage. As I explain elsewhere (J. D. Greene, in press), I believe that the latter usage refers to a natural category of processes involving representations that are inherently neutral, but that may be contingently connected to valenced representations. This allows for the production of behavior that is both flexible and goal-directed.

of mature adults (Kohlberg, 1969), while the other denies that moral reasoning and controlled “cognitive” processes play a direct causal role in shaping ordinary people’s moral decisions (Haidt, 2001).

We conducted a simple behavioral experiment to test our dual-process theory. People responded to dilemmas like the *crying baby* dilemma under normal conditions and under cognitive load (i.e. while simultaneously performing a distracting task that requires cognitive resources). Once again, our claim is that there is an intuitive emotional voice that says “No! Don’t!” as well as a controlled, “cognitive” voice that says, “Please, go ahead.” And if that’s right, then a drain on limited cognitive resources should interfere with the processes that are pushing for “yes,” but not with the more intuitive processes that are pushing for “no.” In the best case, we would expect the imposition of a cognitive load to make people say “no” more often by selectively interfering with the “yes”-friendly processes. And if the load manipulation is not strong enough to actually change people’s judgments, then it might at least make “yes” answers slower, without slowing down the “no” answers. The load manipulation could even *speed up* the “no” answers by knocking out the competition.

This is exactly what we found (J. D. Greene, Morelli, Lowenberg, Nystrom, & Cohen, in prep). When people responded to dilemmas like the *crying baby* case under cognitive load, the load made people slower when they endorsed harming someone in the name of the greater good, but made them a little bit

faster when they said that it would be wrong to cause the harm.² Thus, these results support our dual-process theory of moral judgment: If it were all about intuitive emotional responses, then there would be no reason to think that the cognitive load would slow down “yes” answers any more than “no” answers. And if it were all about controlled processes, then, again, we would have no reason to expect a difference in reaction times between “yes” and “no” answers, as all answers would be slowed by the load. Only a dual-process theory makes sense of this crossover interaction, whereby cognitive load has a different effect on reaction time depending on the content of the judgment.

This experiment tells us that the psychological processes pushing for “no” answers in these cases are rather automatic, charging ahead, impervious to the cognitive load. But this study doesn’t necessarily tell us whether these automatic responses are emotional.³ A different experiment, conducted by Valdesolo and DeSteno (2006) addresses this question. They presented people with two moral dilemmas, which we’ll call the *switch* and *footbridge* dilemmas. In the *switch* dilemma (elsewhere referred to as the *trolley* dilemma (Greene et al., 2001) and the *bystander* dilemma (Thomson, 1986)), a runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save these people is to hit a switch that will turn the trolley onto a side-track

² This facilitation effect was significant only in a sub-group of our participants, namely those who described their decision process as primarily intuitive.

³ It depends on how one defines “emotion.” If a process that is automatic and valenced is necessarily emotional, then this study may be sufficient to implicate emotion.

where it will run over and kill one person instead of five. Is it okay to turn the trolley in order to save five people at the expense of one? Most people say “yes” (Greene et al., 2001; Petrinovich, O'Neill, & Jorgensen, 1993). This case contrasts with the *footbridge* (Thomson, 1986) dilemma: As before, a runaway trolley threatens to kill five people, but this time you are standing next to a large stranger on a footbridge spanning the tracks, in between the oncoming trolley and the five people. The only way to save the five people is to push this stranger off the bridge and onto the tracks below. (You're not big enough to block the trolley by jumping yourself.) He will die as a result, but his body will stop the trolley from reaching the others. Is it okay to save the five people by pushing this stranger to his death? Most people say “no” (Greene et al., 2001; Petrinovich et al., 1993). One might suppose that the action proposed in the *footbridge* case triggers more of an intuitive emotional response than the action proposed in the *switch* case, which would explain why most people tend to say “no” to the *footbridge* case and “yes” to the *switch* case. (Here, too, we're supposing that the controlled process favors the greater good, i.e. the “yes” response.)

Valdesolo and DeSteno tested this hypothesis using an emotion induction paradigm. They presented two groups of people with the *trolley* and *footbridge* dilemmas. The control group, before responding to these dilemmas, watched an emotionally neutral film clip (five minutes from a documentary about a Spanish village). The experimental group watched a five minute comedy clip from *Saturday Night Live*. Valdesolo and DeSteno reasoned as follows: If the typical “no” responses to the *footbridge* dilemma are driven by negative emotional

responses, then hitting people with a dose of positive emotion (using the comedy clip), should counteract the negative emotional response and make those participants more likely to say “yes.” But if the typical “yes” response to the *switch* case is not driven by emotional processes, then watching the clip should have no effect on people’s responses to that case. And, of course, the control group should experience no change as a result of the neutral film. This is what Valdesolo and Desteno found. The neutral film had no effect at all, and the *Saturday Night Live* clip had no effect on people’s responses to the *trolley* dilemma. But the comedy clip did have a significant effect on people’s responses to the *footbridge* dilemma, tripling the number of people who approved of pushing the man off of the footbridge. Thus, it seems that emotional response plays a key role in producing people’s negative judgments to the *footbridge* case.

These two studies nicely support the dual-process model of moral decision-making, providing strong evidence that moral judgment involves both intuitive emotional responses and more controlled “cognitive” processes. Perhaps surprisingly, these two studies were designed to bolster the conclusions drawn from previous neuroimaging studies. In one of these studies (Greene et al., 2001), my collaborators and I presented people with a series of moral dilemmas, including versions of the *switch* and *footbridge* dilemmas. In response to cases like the *footbridge* dilemma⁴, people exhibited relatively higher levels of

⁴ In our initial investigation (J. D. Greene et al., 2001) we used a set of three criteria to distinguish dilemmas like the *footbridge* case (which we called “personal”) from dilemmas like the *switch* case (which we called “impersonal”).

activity in brain regions associated with emotion and social cognition (medial prefrontal cortex, posterior cingulate cortex, and superior temporal sulcus). In contrast, responses to cases like the *switch* case were associated with increased activity in brain regions associated with classically “cognitive” functions such as working memory (dorsolateral prefrontal cortex (DLPFC) and corresponding regions in the parietal lobes). This double dissociation between activity in brain regions associated with emotion, on the one hand, and more classically “cognitive” brain regions, on the other, provided the first piece of evidence in support of our dual-process theory of moral judgment. A second study (Greene et al., 2004) focused on difficult cases like the *crying baby* case. We found that cases like this, relative to easier cases, were associated with increased activity in the anterior cingulate cortex (ACC) and in the DLPFC. Previous work suggests that the ACC plays a role in detecting response conflict (the simultaneous activation of two incompatible behavioral responses; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Botvinick, Braver, Barch, Carter, & Cohen, 2001; MacDonald, Cohen, Stenger, & Carter, 2000), while other work has implicated the DLPFC in cognitive control functions aimed at resolving response conflict

The personal/impersonal distinction was devised as a “first cut” for identifying the psychologically salient features that distinguish these two dilemmas. Based on more recent research (JD Greene, Morelli, Lowenberg, Nysrom, & Cohen, submitted), I now believe that the personal/impersonal distinction should be replaced by a more cognitively sophisticated set of criteria concerning the nature of the agent’s intention and the kind of force applied by the agent.

(Kerns et al., 2004; MacDonald et al., 2000). Thus, increased activity in the ACC and DLPFC is what one would expect if cases like the *crying baby* dilemma create a conflict between two incompatible responses that must be resolved. We also found that when people respond to such cases with a utilitarian judgment (favoring the greater good, even at the cost of harming someone) they exhibited increased activity in the DLPFC. This is what we would expect if the utilitarian “yes” responses are driven by controlled “cognitive” processes.

Like the cognitive load and emotion induction studies described above, these two neuroimaging studies support the dual-process theory of moral judgment, implicating both intuitive emotional responses and controlled “cognitive” responses in moral decision-making. But these neuroimaging studies, compared to the behavioral studies that followed, were a lot more expensive. One has to wonder, then, if they were worth it. Are we getting any additional bang for our neuroscience buck?

Who Needs Social Neuroscience?: The Scientist’s Answer

Neuroscientific data, and neuroimaging data in particular, offer new insights into the relationships among the myriad psychological processes identified using more traditional means. Consider the following remarks from Daniel Gilbert (1999), introducing an edited volume devoted to dual-process theories in social psychology:

The neuroscientist who says that a particular phenomenon is the result of two processes usually means to say something unambiguous—for example, that the inferior cortex does one thing, that the limbic system does another... In such instances the phrase “dual processes” refers to the activities of two different brain regions that may be physically discriminable, and the neuroscientist says there are “two processes” because the neuroscientist is talking about things that can be counted. But few of the psychologists whose chapters appear in this volume would claim that the dual processes in *their* models necessarily correspond to the activity of two distinct brain structures (pg. 3).

Oh, what a decade can bring. Now the social psychologists are also neuroscientists, and they’ve started counting. This transformation, of course, has not yielded precise process counts, and perhaps it never will, given that processes are inherently fuzzy-boundaried things. But the advent of social neuroscience has given students of social cognition, best known for their analytical splitting, the opportunity to engage in an exciting new kind of lumping. By looking directly at the brain we can see whether the processes that Psychologist A has been dutifully teasing apart in her quest to understand Behavior X are, in fact, some of the same processes that Psychologist B has been teasing apart in his quest to understand Behavior Y. Humpty Dumpty may

never fit back together again, but, thanks to neuroscience, he may escape being ground to fine dust.

Thus, neuroscience holds the promise of a new kind of synthetic psychology. It also holds the promise of a different kind of analytic psychology. At present, there is a gulf between the high-level language of the mind (“belief,” “impulse,” “thought,” “attitude,” “emotion,”) and the low-level language of the brain (“lobe,” “gyrus,” “neural activity”). We can correlate things like “attitudes” with things like “neural activity,” but we have only the foggiest picture of how the former arise from the latter. How will we acquire a clearer one? We don’t know how, exactly, or we’d have already done it, but it’s hard to imagine that our learning process won’t involve a fair amount of “top down” investigation, i.e. using what we know about psychology to map the brain in a psychologically meaningful way. An investment in brain-mapping (if done well) will pave the way for a deeper science of mind that is seamlessly integrated with the physical sciences. (But see Uttal (2005))

In short, social neuroscience, at least in the long-run, is likely to yield scientific theories that are richer and more coherent than the ones that social psychologists are used to. This is the standard justification for doing what we’re doing, and it’s a good one. It’s the one that we relay, in various forms, to the editors of journals and administrators of funding agencies, and that is as it should be. But in the long-long-run, the greatest value of social neuroscience may lie elsewhere.

Who Needs Social Neuroscience?: The Philosopher's Answer

When I tell people that I study the neuroscience of moral decision-making, I am often asked, "Where is the brain's moral center?" Apparently, people find the idea of a "moral center" in the brain very compelling. There may be many reasons for this, but I think it has something to do with the fact that a center, unlike a distributed system, need not have *parts*. The moral center of popular conception, I'm guessing, is not a computational system housing an array of dissociable sub-systems that perform relatively simple, complementary functions. It is instead more like a *portal*, out of which fully-formed moral thoughts emerge. A moral portal is what a dualist, when confronted with the fact of the brain, naturally imagines the moral brain to be. The portal theory acknowledges that moral judgments must get into our brains somehow, while leaving open the possibility that their true origin lies beyond.

This dualist conception of moral judgment breaks down when the moral brain is functionally decomposed. This involves two things. First, the process of moral decision-making is itself broken down into distinct psychological processes. Second, it is shown that distinct parts of the brain are respectively responsible for carrying out these distinct processes. Both aspects of functional decomposition are necessary if they are to count against dualism. If the component processes are distinguished psychologically, but not attributed to different parts of the brain, then we are free to think of these psychological

processes as operations of a multi-faceted soul that renders its judgments before transmitting them to the material realm via the brain's moral portal. Consider, for example, the cognitive load and emotion induction studies described above. The cognitive load study tells us that, when confronted with cases like the *crying baby* dilemma, we have an intuitive response that tells us one thing ("Don't smother the baby!") and a more controlled response that tells us the opposite ("But there is nothing to lose and much to gain by smothering the baby.") Valdesolo and DeSteno's (2006) emotion induction study teaches a complementary lesson, demonstrating that our judgments, in some cases more than others, are driven by emotional responses. These results provide evidence for the dual-process model of moral judgment, but they do little, if anything, to dispel dualism. Dualists and materialists alike are familiar with having emotional impulses, and with resisting them. A dualist can happily regard these mental operations as operations of the soul, and these behavioral experiments provide no evidence to the contrary.

The situation changes little if we point to multiple brain regions that are "involved" in moral judgment, but say nothing about how these various brain regions contribute to the decision-making process. We can stick someone in a brain scanner, have that person make moral judgments, and then report on the brain regions that exhibit increased activation, but this will do nothing to embarrass a dualist. He can happily regard this smattering of brain regions as a distributed portal, an archipelago of mysterious mind-body interaction.

It is only when we ascribe distinct psychological sub-functions of moral decision-making to distinct physical parts of the brain that moral decision-making starts to look like a mechanical process. And this is exactly what the neuroimaging experiments described above do. Our claim is not that the ACC is “involved” in moral judgment in some nebulous way. If our theory (Greene et al., 2004) is correct, the ACC performs the specific function of detecting conflict between competing responses, both in moral contexts and in other contexts that have nothing to do with morality *per se* (Botvinick et al., 2001; Cohen, 2005). Our theory likewise attributes a control function to the DLPFC and an emotional function to the medial prefrontal cortex⁵. Of course, the dualist can dig in his heels, even in the face of functional decomposition of this kind. He might imagine, for example, that each of these brain regions simply receives transmissions from different functional parts of the soul: The ACC receives transmissions from the part of the soul that detects conflict, and so on. But this

⁵ More recent neuropsychological data provide additional support for our conclusion that parts of the medial prefrontal cortex generates or mediates emotional responses that drive non-utilitarian judgments in response to moral dilemmas such as the *footbridge* case. Mendez *et al.* (Mendez et al., 2005) found that patients with frontotemporal dementia (known for their “emotional blunting” and ventromedial prefrontal damage) were disproportionately likely to approve of pushing the man off of the *footbridge*. Similar results have been obtained in patients with more well-defined ventromedial prefrontal lesions (Koenigs et al., submitted).

can't go on forever. As our body of knowledge grows, the moral brain will be decomposed into smaller and smaller physical parts, associated with narrower and narrower psychological functions, until the corresponding bits of soul are reduced to an array of manifestly superfluous micro-ghosts with no purpose other than to be there.

As Wolfe's remarks suggest, the soul will officially expire when the mechanics of the moral mind become transparent. I believe that the death of the soul may prove to be one of psychology and neuroscience's most lasting contributions. That is, if we're around long enough to get the job done.

Dualism: What's at Stake?

Why does it matter if people are dualists? As long as we scientists know what we need to know in order to do our work, is it our business, or in anyone's best interest, to provide compelling demonstrations of the fact that we have no souls? I think that it is. Dualist beliefs may be harmless enough most of the time, but they divide us in destructive ways, enable us to do some of the worst things that we do, and may ultimately lead to our demise.

Consider, once again, the events of 9-11. Nineteen men killed nearly three thousand people, setting in motion a series of events that, in addition to killing many thousands more people, has destabilized the Middle East at a time when nuclear weapons are becoming increasingly accessible. Those nineteen

men destroyed their bodies, along with thousands of others, believing that they—their souls—would go on to enjoy a pleasant post-corporeal existence. Of course, it's possible that their beliefs about the next world played no role in their decision to leave this one, but that seems unlikely. Rather, as others have noted (Dawkins, 2006; Harris, 2004), it seems that their beliefs about the afterlife enabled them to do what they did.

Here in the West, our leaders are dualists, too. George Bush does not speak openly about the details of his metaphysical worldview, but as an evangelical Christian he presumably believes that all people have souls, and that the souls of those who share his faith will be saved, while the remainder will spend eternity in hell. This remainder presumably includes the vast majority of Muslims. Thus, transforming a Muslim society into a more Westernized society with greater exposure to Christianity may translate into huge gains in terms of the number of souls saved. The prospect of saving millions of souls may warrant taking large risks. More specifically, it might warrant risking the lives of millions of Muslims, assuming that their souls are, in the absence of Christian intervention, doomed. According to one report, over half a million Iraqis have died as a result of the American invasion of Iraq (Tavernise & McNeil, 2006).

Dualism is at the heart of many bioethical debates (Bloom, 2005). It is often said that the abortion controversy is really a debate about when life begins. But “life” is not the real issue, as everyone agrees that a fertilized human egg, like an unfertilized human egg, is alive. Nor is it a matter of when a developing human acquires significant psychological characteristics such as the ability to

feel pain. Most opponents of abortion are perfectly happy to eat animals that are capable of feeling pain, etc. Nor is it a matter of destroying potential human life. Both birth control and abstinence rob potential humans of their existence. Rather, the debate over abortion is ultimately a metaphysical one. The question is not whether a fertilized egg is alive, but whether it is host to a “human life,” i.e. a human soul. Without a soul in the balance, there is no abortion debate. Likewise for the debates over human stem cell research and euthanasia.

As I (Greene & Cohen, 2004) and others (Bloom, 2005) have argued, certain aspects of our criminal justice system are implicitly dualist. If you ask people why we ought to punish criminals, people most often cite the law’s deterrent effect. But when people respond to concrete cases, their judgments are surprisingly insensitive to factors that are relevant to the prevention of future crime (Carlsmith, Darley, & Robinson, 2002). Rather, it seems that people’s intuitions about punishment are *retributivist*. We want to punish criminals, not because of the future benefits, but simply as an end in itself. These retributivist tendencies are, I believe, implicitly dualist. If someone has a brain tumor that causes aggressive behavior, people are far more willing to forgive that person. “After all,” we say, “It’s not *him*, it’s *his brain*.” When we attribute bad behavior to a purely physical cause (such as a brain tumor), the retributivist impulse fades. Our aim is to punish guilty minds (*mens rea*), not broken brains. (A broken brain may be worth containing, deterring, and rehabilitating, but there’s no good reason to punish someone simply for having a broken brain.)

From a neuroscientific perspective, of course, all behavior (good and bad) has purely physical causes, and anyone who does unusually bad things must have something, however subtle, wrong with his brain. Combine this ordinary scientific assumption (all bad behavior is caused by brains that are, in some sense, broken) with people's ordinary assumption about punishment (there is no inherent value in punishing someone for having a broken brain), and we get a very different sort of legal system. We get one focused exclusively on the practical business of preventing future crime, rather than on the metaphysical business of making guilty minds suffer for their sins. In the United States, at least, our prison system is very good at making people suffer, but its merits as a system for preventing future crime are highly questionable (Tonry, 2004). If we were more interested in reducing crime, and less interested in making guilty minds suffer, we might all be better off.

Dualism plays a parallel role in people's thinking about mental illness. Intuitively, we all agree that people with cancer deserve our sympathy and financial support because cancer is a serious medical problem. But if someone is depressed, that person's condition is, to many people at least, just "psychological," and the prescription is to "snap out of it." Dualism draws an illusory distinction between having a weakened body and having a weakened mind.

Finally, dualism may play an important role in people's attitudes concerning the environment. According to a variety of polls, about 40% of Americans believe that we are living in the "end times," i.e. that the world will end

relatively soon, at which point all followers of the Christian faith will be swept up into heaven, while the rest of us descend into hell (Sahagun, 2006). If you think that God is going to end the world relatively soon, you're unlikely to be terribly concerned about the level of carbon dioxide in the Earth's atmosphere. Of course, most of the people who are worried about global warming are dualists, too. But to be indifferent to the fate of this world, it certainly helps to believe that there's another one waiting.

Conclusion

Social neuroscience is exciting, but it's hard for some of us to say why. Most would agree that looking directly into the human brain will, sooner or later, provide us with better theories about how our minds work. But the prospect of better psychological theories, arriving sooner or later, hardly explains the excitement we feel. I believe that social neuroscience is exciting primarily because of its broader philosophical implications, and only secondarily because of the empirical details we expect it to yield. But to speak of social neuroscience's philosophical "implications" is a bit awkward. Officially, we scientists already know that the operations of the mind are the operations of the brain, and not those of an immaterial soul. This is, at the very least, our working assumption. In making this assumption, however, we part ways with the rest of humanity, the vast majority of whom explicitly believe that we are souls housed in

bodies. Such dualist tendencies are, in my opinion, a major social problem, and may become increasingly destructive. If that is correct, then dispelling dualism is serious business, at least as serious as curing cancer, and probably more so. If anything can cure us of our dualist tendencies, it is social neuroscience, the physical science of human experience. By decomposing the social brain into its mechanical components we can do good science in the conventional sense, but that is, I think, the least of what we're doing. Social neuroscience is, above all else, the construction of a metaphysical mirror that will allow us to see ourselves for what we are and, perhaps, change our ways for the better.

Acknowledgments: Many thanks to Andrea Heberlein and Joe Paxton for helpful suggestions.

- Bloom, P. (2004). *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books.
- Bloom, P. (2005). Worse than creationism. *American Psychological Society Observer*, 18(10).
- Bloom, P. (2006, June, 27, 2006). Seduced by the flickering lights of the brain. *Seed*.
- Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, 402(6758), 179-181.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev*, 108(3), 624-652.
- Cacioppo, J., Bernston, G., Lorig, T., Norris, C., Rickett, E., & Nusbaum, H. (2003). Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles. *Journal of Personality and Social Psychology*, 85(4), 650-661.
- Carlsmith, K. M., Darley, J. M., & Robinson, P. H. (2002). Why do we punish? Deterrence and just deserts as motives for punishment. *J Pers Soc Psychol*, 83(2), 284-299.
- Cohen, J. (2005). The vulcanization of the human brain: A neural perspective on interactions between cognition and emotion. *Journal of Economic Perspectives*, 19, 3-24.

- Dawkins, R. (2006). *The God Delusion*. Boston, MA: Houghton Mifflin.
- Dennett, D. C. (2006). *Breaking the Spell: Religion as a Natural Phenomenon*. New York: Viking.
- Gilbert, D. T. (1999). What the mind's not. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology*. New York: Guilford.
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philos Trans R Soc Lond B Biol Sci*, 359(1451), 1775-1785.
- Greene, J., Morelli, S., Lowenberg, K., Nysrom, L., & Cohen, J. (submitted). *Cognitive load selectively interferes with utilitarian moral judgment*. Unpublished manuscript.
- Greene, J. D. (in press). The Secret Joke of Kant's Soul. In W. Sinnott-Armstrong (Ed.), *Moral Psychology: Morality in the Brain* (Vol. 3). Cambridge, MA: MIT Press.
- Greene, J. D., Morelli, S., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (in prep). *Cognitive load interferes with utilitarian moral judgment and facilitates non-utilitarian moral judgment*. Unpublished manuscript.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2), 389-400.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105-2108.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834.
- Harris, S. (2004). *The End of Faith: Religion, Terror, and the Future of Reason*. New York: W. W. Norton & Co.
- Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am Psychol*, 58(9), 697-720.
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., 3rd, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303(5660), 1023-1026.
- Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., et al. (submitted). Damage to the prefrontal cortex increases utilitarian moral judgments.
- Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago: Rand McNally.
- Lieberman, M. D., Gaunt, R., Gilbert, D. T., & Trope, Y. (2002). Reflection and reflexion: A social cognitive neuroscience approach to attributional inference. *Advances in Experimental Social Psychology*, 34, 199-249.
- MacDonald, A. W., 3rd, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472), 1835-1838.
- Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An investigation of moral judgement in frontotemporal dementia. *Cogn Behav Neurol*, 18(4), 193-197.

- Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology*, 64, 467-478.
- Sahagun, L. (2006, June 22, 2006). 'End Times' Religious Groups Want Apocalypse Soon. *Los Angeles Times*.
- Tavernise, S., & McNeil, D. (2006, October 11, 2006). Iraqi dead may total 600,000, study says. *New York Times*.
- Thomson, J. J. (1986). *Rights, restitution, and risk : essays, in moral theory*. Cambridge, Mass.: Harvard University Press.
- Tonry, M. (2004). *Thinking about crime: sense and sensibility in American penal culture*. New York: Oxford University Press.
- Uttal, W. R. (2003). *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain*. Cambridge, MA: MIT Press.
- Uttal, W. R. (2005). *Neural Theories of Mind: Why the Mind-Brain Problem May Never Be Solved*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Valdesolo, P., & DeSteno, D. (2006). Manipulations of emotional context shape moral judgment. *Psychol Sci*, 17(6), 476-477.
- Wegner, D., & Gilbert, D. T. (2000). Social psychology--the science of human experience. In H. Bless & J. P. Forgas (Eds.), *Subjective Experience in Social Cognition and Behavior*. Philadelphia, PA: Psychology Press.