Self is Magic

Daniel M. Wegner

Harvard University

Chapter prepared for J. Baer, J. C. Kaufman, & R. F. Baumeister (Eds.),

Psychology and Free Will. New York: Oxford University Press. (12-27-2006).

Imagine a magician who can make things happen merely by *thinking* of them. This magician thinks "I'd like the lights on," and before you know it…there is light! Right there and then, a hand has reached out and turned on a lamp. Next, the magician hopes for warmth. In a matter of minutes, a fire is glowing and crackling in the fireplace. The magician's wish for a cup of tea and a plate of cookies soon unfolds into just the right tea and just the right cookies, placed conveniently at arm's length. Every wish seems to come true. And when the next cookie comes to mind, the arm reaches to the plate, grasps a cookie (not just any cookie, but the very one that the magician is thinking about) and deftly brings it to the conjurer's mouth at the appropriate angle for a nice bite. What wouldn't we all give to live such an enchanted life?

Aside from the occasional stray cookie that slips to the floor, we do live this life. Human action is a kind of magic, an astonishing ability to think of something and thereby make it happen. Although a lamp will seldom light on its own merely because we want it to, we find that our fingers leap to the switch and light that lamp when the idea comes to mind. Our actions are an astonishing realm of events that bend to our desires when so much of the world does not. Perhaps this is why each person views self with awe—*The Great Selfini* amazes and delights! We are enchanted by the operation of our minds and bodies into believing that we are "uncaused causes," the origins of our own behavior. Each self is magic in its own mind.

Unfortunately, the magic self stands squarely in the way of the scientific understanding of the psychological, neural, and social origins of our behavior and thought. As long as we are charmed by ourselves, perhaps we won't see beyond the magic. This chapter examines this curious standoff in several ways. We will begin by

exploring just how it is that the self's magic seems to arise. We will try to "reverse engineer" the magic, discerning what kind of system might be necessary to make us into such impressive beings that we are spellbound by our own performances. Then, we will consider why it is that this concept of self as an inner origin of our actions is so resistant to analysis and understanding—why even those of us who hope to overcome our belief in magic are still captivated every day by the self's parlor tricks. And finally, we will turn to the question of why the process of evolution might have unfolded in such a way as to make us think we are magical creatures.

Making Magic: The Birth of an Ego

How would one go about making a being that believed in its own magic? Could a robot or computer be constructed with this property? This entity would need to have many of the standard cognitive properties of the human. Of course, it would need to be able to perceive events and develop causal theories of their interrelations, and it would need to have the ability to reflect on itself and focus those causal theories on its own processes. These are standard features that are often discussed by those who imagine reverse-engineering a human mind (e.g., Angel, 1989; Braitenberg, 1984; Dennett, 1994; Holland, 2003; Scassellati, 2002). But to believe itself magical, this entity would need to have one property we don't often appreciate in the cognitive toolkit of the standard human: It would need to have *incomplete self-knowledge*. Perceiving magic of any kind requires that we don't fully understand how something has happened.

This realization follows directly from the theory of magic perception introduced (with a flourish) by Harold H. Kelley (1980). According to Kelley's analysis, we perceive magic when an apparent causal sequence shortcuts or obscures a real causal

sequence that is not itself fully perceived (and the apparent sequence usually departs from one that common sense would predict). When the magician saws a lady in half, for example, an event has occurred that is only apparent: the lower half of her body seems to have seceded from the upper half. We know this can't be the real causal sequence or there would be a lot of magician's assistants out there whose careers were tragically cut short. There *is* a real causal sequence that may involve contortion, mirrors, special cavities in the box, an additional lady to provide distal wiggling feet, and what else only the magician knows. But the audience sees the apparent causal sequence in which the saw seems to cut her in half, and if the illusion is well done, they are amazed by the trick. It is in exactly this sense that the self is magic: When we look at ourselves, we perceive a simple and often astonishing apparent causal sequence (I thought of it and it happened!) when the real causal sequence underlying our behavior is complex, multi-threaded, and unknown to us as it happens.

Now, saying that people believe in magic, at least in our enlightened times, can be something of an insult. The film *Jesus is Magic* by comedienne Sarah Silverman, for example, chides Christians about their beliefs and suggests, all in jest of course, that they are being duped somehow. Belief in magic may be fun and even a source of delight, but it also can entail childlike naiveté, delusion, or just plain foolishness. Seeing one's own causal influence as supernatural is part of being human, though, so rather than ruing this human tendency or calling it foolish, it is psychological science's job to understand it. How do people develop this magic self—what Dennett (this volume) calls "some concentrated internal lump of specialness"? Why do we experience our actions as freely willed, arising mysteriously from the self, and why too do we resist attempts to explain

those actions in terms of real causal sequences, events that are going on behind the curtain of our minds?

One explanation of the magic draws on the idea that the mind only presents us with a relatively impoverished account of its own operations, and our attempt to make sense of the evidence yields the impression that we are freely willing our actions. This account is the basis of the theory of *apparent mental causation*, a set of ideas that draw on the philosophy of Hume (1739/1888) to explain how it is that people come to experience conscious will (Wegner, 2002, 2003, 2004; Wegner & Wheatley, 1999). This theory and several related accounts (e.g., Brown, 1989; Claxton, 1999; Michotte, 1963; Nisbett & Wilson, 1977; Spence, 1996; Thompson, Armstrong, & Thomas, 1998) propose that people experience willing their actions when they draw causal inferences relating their thought to their action. Quite simply, a person infers that an event is due to the self as a result of perceiving a causal link between own thoughts and that event. Hume's insight was to note that such perception is a matter of inference, not direct perception, and the implication of this insight is that the perception of one's own causality is open to error.

How do we go about drawing this causal inference about our own action? Consider that magical act of turning on the light. This is something that sometimes can feel quite willful, and at other times can feel absent-mindedly automatic. If you have just thought about turning on the light and then do so, it may feel more willful—whereas if you have been thinking about having a cookie and then suddenly find yourself turning on a light instead, it is likely to feel less willed and more like some sort of alien control. To support a feeling of will, the thought of turning on the light also must occur just prior to

the action to maximize the experience of will, as thoughts that occur far beforehand (and that then are forgotten until the action), or thoughts of flipping that switch that only appear after the light is on do not seem to prompt a sense of willed action. And if someone else presses your hand to the lamp, you may discount entirely the causal role of your prior thought and feel the act is unwilled. These observations point to three key sources of the experience of conscious will–the *consistency*, *priority*, and *exclusivity* of the thought about the action. For the perception of apparent mental causation, the thought should be consistent with the action, occur just before the action, and not be accompanied by other potential causes. Several studies have examined the influence of each of these principles of apparent mental causation.

Consistency effects. The idea that a thought and action must be consistent with each other to yield an experience of conscious will has been investigated in experiments on magical thinking (Pronin, Wegner, McCarthy, & Rodriguez, 2006). These studies examined whether priming people to experience thoughts consistent with events they did not actually cause might lead them to experience the events as caused by the self.

Participants in one experiment were asked to play the role of a witch doctor in a study of psychosomatic influences on health, and to perform a voodoo curse by sticking pins in a doll in the presence of another participant assigned to play the role of the victim. The victim role was in fact played by an experimental confederate, who later feigned a headache. The question of interest for this research was whether participants would accept any causal responsibility for this headache: Would they believe that they had exerted some influence over the victim's health? A participant would not need to believe in voodoo per se to reach this conclusion, but would only need to perceive that a potential

victim might be stressed into a headache by the shock of receiving a curse. Participants were given an article suggesting this possibility in preparation for the experiment (Cannon, 1942), and many did accept the idea that their actions caused the headache.

Magical thinking was amplified in this experiment by a manipulation of the consistency of participants' thought with the action. The inference that "I made the victim sick" was significantly strengthened among participants who were led to dislike the confederate before performing the voodoo curse. These participants were exposed to a confederate who was late, rude, and messy; post-experimental questioning revealed that they had indeed come to dislike this person. As compared to participants who met an unremarkable confederate—one who was normal and likable—those who performed the curse on the confederate victim they disliked were more inclined to believe that their curse had caused the victim's headache. In a follow-up study, the same phenomenon was observed when participants were merely instructed to "think negative thoughts" about the victim before pinning the doll. As compared to those who were not given this instruction, the participants led to think in a way that was consistent with the act of harming the victim came to believe that they had indeed caused such harm. And of course, no harm had been caused at all.

These findings suggest that people can easily develop the belief that they are harming someone when they have wished for such harm, even when the harm befalls the victim for reasons unrelated to the harm-wisher's desires. This may be why we can feel guilty if we have wished ill on someone and they suffer an unfortunate fate—even though we know we really were not responsible. Cursing an obstreperous elderly relative under our breath may lead us to feel particularly culpable when the relative falls and breaks a

hip. Our natural tendency to link our thoughts with consistent events leads us all too readily to the mistaken belief that the events have issued from our will.

The self can also be magic in a good way. To study this possibility, we put the voodoo doll in mothballs and turned to the magic of cheerleading. Further studies by Pronin et al. (2006) revealed that there is a tendency to believe that one has caused a positive event to occur merely because one has imagined it. In these experiments, people were led to think about the outcomes of sporting events being played by others, but were questioned afterwards about their own causal influence on the sporting outcomes. People who were asked to envision the success of a basketball shooter on each of 8 free-throws were more inclined to believe they had indeed helped him than those who were asked to envision him lifting barbells when he then proceeded to sink 6 of the 8 shots. People who were watching all this and were informed about what the spectators were visualizing reached the same conclusion: Even these uninvolved onlookers thought the spectator visualizing successful shots had somehow helped the shooter to succeed.

This tendency to claim authorship for the successes of others also extends to one's favorite sports teams in real competition. In another study, fans at a basketball game were asked to complete a pre-game exercise—either thinking about how each of their team's players could contribute to the game, or thinking about how each player could be identified in a crowd. When the fans were then quizzed in the middle of the game to see whether they felt they were personally influencing the game's outcome, those who had been prompted to develop success-relevant thoughts for their team were more likely to report exerting influence. A final study by Pronin et al. (2006) found that fans watching the 39th Super Bowl football game on television were susceptible to the same illusion.

Those viewers who reported thinking more about the outcome of the game also claimed more personal responsibility for the game's outcome—regardless of whether their favorite team won or lost (the Patriots won, by the way, largely because I had wished this).

These studies highlight what may be a general process of mind underlying belief in paranormal phenomena such as ESP, clairvoyance, precognition, and psychokinesis. In everyday life, our bodies appear to respond readily and easily to many of our wishes. Yes, we may find it difficult to wish to perform a Chopin etude on the piano when we've never taken a lesson—but there are so many things we *can* do, things that happen just because we want them. It makes sense that this normal human capacity for conscious will might lead us into overextensions from time to time. If our wishes seem to prompt a range of activity within our personal sphere of influence, why not hope for more? The many forms of supernatural belief, including beliefs in the effectiveness of appeals to deities, may develop as natural next steps that follow from the magic we perceive in ourselves. If mere wishing can pop the lid off a bottle of beer, why not wish for the moon?

The belief that one is influencing events also can be enhanced by prior action-consistent thoughts that are not conscious. Pre-action subliminal primes of action effects can increase perceptions of authorship for the action. This result was observed in a study when people were asked to judge whether their button press was responsible for the resting position of a marker on a computer display (Aarts, Custers, & Wegner, 2004). There were two markers moving very quickly on the display, and the participant's task was to judge whether their marker was the one that had come to that resting spot. On

some trials, the resting spot was primed with a brief flash at that position, and it was on these trials that participants estimated more often that the marker was their own—even when the flash priming that position was so brief as to be subliminal. In subsequent research, this effect has also been observed among depressed individuals—suggesting that the consistency of thought and action enhances perceptions of own agency even among people with weakened self-views that surface in depression (Aarts, Wegner, & Dijksterhuis, 2006).

The consistency of thought and action can be undermined even in normal action when people are distracted from their action-consistent thoughts. People in one study were asked to suppress thinking about what they were doing as they performed each of a series of simple activities—such as winding thread on a spool (Wegner & Erksine, 2003). They reported being somewhat successful at the suppression, and also reported weakened feelings of voluntariness for these actions. The active disengagement from thoughts about actions may be a pathway leading to the phenomena of hypnosis, a kind of "voluntary involuntariness" that comes about when people are instructed to ignore their own thoughts about what they are doing (Lynn, Rhue, & Weekes, 1990; Lynn, Weekes, Matyi, & Neufeld, 1988).

Priority effects. The second principle of apparent mental causation, that self will be seen as causal when the thought of action occurs just prior to the action, has also been tested in research. Priority effects were observed initially by Wegner and Wheatley (1999), and have also been obtained in studies of vicarious agency by Wegner, Sparrow, and Winerman (2004). For these latter experiments, participants were led to experience the arm movements of another person as if the movements were their own. The

participant was attired in a robe and positioned in front of a mirror such that the arms of a second person standing behind the participant could be extended through the robe to look as though they were the arms of the participant. The second person wore gloves to aid in this illusion. Participants kept their own arms at their sides and were instructed not to move. Both participant and "hand helper" wore headphones.

For the experiments, the helper's arms performed a series of 32 movements (e.g., snap the fingers of your right hand, wave hello with both hands, etc.) in response to sequential instructions the helper was given via the headphones. In one experiment, participants also heard the instructions for each of the arm movements through their own headphones, or they heard nothing. Those who heard the instructions thus were provided with consistent prior thoughts for actions they perceived visually to be occurring in the position their own actions might occur. As might be expected from the aforementioned consistency studies, the consistent previews led participants to report enhanced feeling of control over the arm movements as compared with other participants. Participants did not feel that they had full control of the arms, of course, as they had no control at all—but they reported a significantly enhanced *impression* of such control.

Another study in this paradigm tested the effects of priority. For this study, instructions were given a few seconds before each movement, just prior to movement, or after each movement had occurred. Participants felt decreased control over the arms' motions with late instructions, whereas the slightly early and just-in-time instructions yielded similarly enhanced experiences of agency as compared to a no-instruction group. These results suggest that even a minor shift in timing—in which an instruction appears just a moment too late—irremediably undermines the illusion of agency that the

instruction provides. Knowing what another person's arms are doing after they've finished their motion produces no notable increment in the feeling of vicarious agency for that motion.

The priority effect seems well illustrated in the feeling of uncanny agency that comes when we serendipitously anticipate an event. Thinking about a friend just before the friend calls on the phone, for example, prompts an odd sense of agency—we feel as though we'd conjured them up (Blackmore, Galaud, & Walker, 1994). Experiences like this one are so profound that they often are reported as the front line of evidence among believers in supernatural phenomena ("ESP must exist—remember that time I thought of Aunt Milly from Idaho just before she called?"). Thinking of the caller afterwards, of course, would be entirely unimpressive, as the absence of proper priority would undermine any sense of personal agency. The matter of timing is crucial in the perception of willed action, so crucial that even happenstance events may be perceived as under one's control when they occur just after one has happened to think about them.

Exclusivity effects. The third principle of apparent mental causation is that people see their thoughts as causing events to the degree that there are no other plausible candidate causes. When the thought and only the thought precedes an event—no one else is thinking something similar, for example, or doing something that seems influential—the person will experience the event as flowing from that thought.

We are often quite sensitive to the possibility that there are causes beyond our own thoughts that might produce our action. When we are thinking of having the halibut in a restaurant, for example, and someone else at our table orders the halibut just before we've been able to say that's what we want, we can feel "scooped." We may think that

we should pick something else so it doesn't look as though we were merely copying our fellow diner. This sensitivity suggests a more general readiness to perceive that our conscious will is challenged by external authors. We become ill at ease when we are faced with questions of our own free will not so much because we have some aversion to causal determinism arising within us, but because of the concern that we are being pushed around or influenced by others (Bargh, this volume; Wegner & Sparrow, 2004). It is curious that we humans have bodies and minds so well constructed for proprioception—we can perceive our actions through intricate pathways of muscle sense, vision, joint and tendon movement sensations, vestibular senses, and more (e.g., Jones, 1988)—but we nonetheless discount all this internal evidence of our own causal influence when other people might be causing our actions. All it seems to take is another's prior movement, or even the hint of a command, and we relinquish much of our own experience of will and allocate responsibility to the other.

A renowned instance of this effect occurred in the obedience studies conducted by Milgram (1963). Research participants were led to believe that they were teaching another participant in an experiment by applying electrical shocks whenever he performed incorrectly, and many were found to apply such shocks willingly—to the point of apparently placing him in grave danger and possibly causing his death. Yet these people were only willing to accept a modicum of responsibility for this action. Participants obeying the experimenter reported what Milgram called an *agentic shift*: "the person entering an authority system no longer views himself as acting out of his own purposes but rather comes to see himself as an agent for executing the wishes of another person" (Milgram, 1974, p. 133).

Exclusivity effects on the experience of will have been observed in studies of sensitivity to fine differences in the timing of action and gaze (Sparrow & Wegner, 2006). For these studies, a participant was asked to tap out the letters of the alphabet in order with a conductor's baton by following a line connecting letters on a maze, each on the click of a metronome. This was entirely straightforward for participants, as all were fairly familiar with the alphabet. After each completion of the maze, participants took a minute to rate the action on a set of scales measuring their experience of authorship. The study found that the experience of will was reduced when the experimenter pointed or merely gazed at the alphabet letters one in advance of the participant's current letter—and that the experience of will was enhanced when the experimenter pointed or gazed at the letters one behind the participant's current letter. Apparently, it doesn't take much in the way of social circumstances to override the sources of authorship information in body and mind to produce an alteration in overall experience of will. Someone else doing an action just before or after we do it makes us feel differently about whether we did it.

Drawing causal inferences about our own thoughts, then, is a major way in which we develop experiences of willing what we do. The sense of magic in the self is produced by mental processes that perceive the consciously accessible parts of the action puzzle—the thoughts about the action that come to mind, and the perceptions of the action itself. Sometimes these processes have access to yet other sources of evidence, as when people use perceptions of their own effort to draw inferences about their authorship of action (Preston & Wegner, in press). But experience of apparent mental causation renders the self magical because it does not draw on all the evidence. We don't have access to the myriad neural, cognitive, dispositional, biological, or social causes that have

contributed to the action—nor do we have access to the similar array of causes that underlie the production of the thoughts we have about the action. Instead, we look at the two items our magic selves render visible to us—our conscious thought and our conscious perception of our act—and believe that these are magically connected by our will. In making this link, we take a mental leap over the demonstrable power of the unconscious to guide action (e.g., Bargh, 2004; Bargh & Barndollar, 1995; Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trotschel, 2001), and conclude that the conscious mind is the sole player. We discern our magical role in the world by reference to any apparent premeditation of the world's events that we can find in our own minds.

Breaking the Spell: Trying to Look Behind the Curtain

Does knowing how the trick works undo the magic? Many of the most strident arguments for free will hinge on the idea that a scientific understanding of human behavior could potentially ruin everything. The magic will be undone, the glorious human spirit will be cheapened, demystified, and rendered grotesque. We will uncover the trolls operating the machinery in the dungeon, and we will never again be able to appreciate the sparkling radiance of the Magic Kingdom of the self. Or more realistically: we will uncover the genetic codes that produce neural structures that allow incoming sensations produced by social and situational factors to contribute to the cognitive computations that incline our motor output processes to lead us to behave—and *then* we lose the magic. Now of course, we tend to worry about the dangers of unweaving the rainbow in all of science (Dawkins, 1998)—but we most seem to fear the loss of magic that might ensue if we came to understand ourselves (Dennett, 1984, 2003;

Greene & Cohen, 2004). When we know the trick of what makes us behave, the magic of free will could be lost.

Magic does seem a fragile thing in some cases. Consider the magic of love. When people in close relationships are asked to explain why they feel love toward their partner, they subsequently rate their love for their partner as less than if they were prompted to give no explanation at all (Seligman, Fazio, & Zanna, 1980). The love doesn't go away entirely, but it does decline a bit for the moment. This same effect occurs for other emotional states; a number of studies have revealed that explaining a feeling can have the effect of dissipating that feeling (Pennebaker, 1997; Wilson, Gilbert, & Centerbar, 2003). There is something about portraying deep and meaningful things in layers of analysis that makes their meaning less compelling. Just try to explain a joke to someone and watch the mirth drain from the room.

Explanations also seem to dismantle the magic of evil. One of the key fears of psychologists who try to understand and explain the behavior of criminals, psychopaths, or other villains, is that their explanations render the evil person into someone normal who is just behaving in a world that seems to require evil. Explaining evil, in other words, seems tantamount to condoning it (Miller, Gordon, & Buddie, 1999). To retain our full appreciation of evil—enough at any rate to work up an appropriate level of outrage and hatred for the evildoer—we almost need to *resist* understanding it. Roy Baumeister's (1997) superb book on the psychology of evil returns to this theme repeatedly, as he reveals the conflict we all feel in trying to establish a causal understanding of evil without simultaneously making it somehow less horrible and repugnant.

Generating explanations for the existence of our beliefs may also reduce their perceived value. Preston and Epley (2005) have found that when people give explanations for a belief, they feel the belief is less valuable. When people use the belief as an explanation for other things, however, they feel that the belief is more valuable. Certain beliefs, such as a belief in God or a belief in personal free will, are easy to use as explanations—we could spend all day talking about what God could do, or what a person with free will might choose. But these sorts of beliefs are themselves difficult to explain. Why does God do what He does, and why does free will opt for fried dumplings instead of mixed greens? The finding that people who are urged to explain beliefs feel the beliefs are less valuable, like the findings for love and for evil, reveal that the simple process of explaining may indeed dispel magic.

These studies suggest that it is possible to "break the spell"—given the right spell and the right explanatory counter-spell. There do seem to be cases when thinking about how something works can reduce our experience of magic. But is this the natural result of psychological explanation? Must the self be destroyed by its own explanation? The fact is, there are many mysteries that do *not* lose their poetry merely because they have been solved. Do people actually fall out of love and part forever when they've paused to "count the ways"? Probably not. And understanding evil doesn't make us treat it much differently either. Many legal cases have been tried in which those accused of crimes have attempted to escape punishment with a "good explanation"—usually an insanity plea of some kind (Denno, 2002)—but aside from a few anomalies, these explanations typically do not shield the accused from punishment. Indeed, the forms of incarceration

that have been invented for people who plead insanity are sometimes more odious than those for people who are merely found guilty (Monahan et al., 1995).

The magic of self does not readily go away when we explain action. In fact, the self seems remarkably resistant to reports of its demise, cropping up again and again in most every living human. One reason for this resiliency could be that the self is an ongoing experience (Wegner, 2004). Every time we think of doing something and then do it, we play the trick on our minds. It seems like magic each time. Even if we have somehow overcome the magic we experienced when we thought of getting a cup of coffee and found it in our hands—the next moment we are thinking of going to the window to check on the weather and there is magic once again! Perhaps the sheer frequency of our experiences of conscious agency is sufficient to overwhelm the nattering of our inner skeptic telling us our behaviors are caused by mechanisms of mind and not by our free willings. We could be convinced of the magic self by its mere doggedness.

But there is more. Just as a joke that is repeated again and again becomes less funny, there is probably nothing about the frequency of free willing alone that keeps the magic of self alive. Repetition doesn't build the illusion. Rather, the illusion of the magic self is *inherently persistent*. This is a trick that we can't see through, an illusion that cannot be spoiled by knowledge of how and why it happens. There are visual illusions that have this power—they continue to fool us even though we know they're illusions. For example, Roger Shepard (1990) drew tables that have this property (see Figure 1). The table on the left looks longer than the one on the right—in fact, the two of them *seem* very different. Yet if you cut out one table top and lay it on the other, you will find that the surfaces coincide exactly. You probably shouldn't really cut them out or it

will ruin this perfectly nice book—just take my word for it (and Shepard's), and truly believe in your heart that the two surfaces are identical. And despite this belief, you will find that each time you look at them, the one on the left looks longer.

- - - - - - - - - - -
Insert Figure 1
- - - - - - - - - - -

The magic of self is just like this. It is not logical. It doesn't go away when you know how it works. It still feels as though you are doing things, freely willing them, no matter how much you study the mechanisms of your own behavior or gain psychological insight into how all people's behavior is caused. The illusion of self persists. This is why hand-wringing about the potential dangers of determinist thinking for morality (Shariff, Schooler, & Vohs, this volume) or for the law (Greene & Cohen, 2004) doesn't make a whole lot of sense. The agent self is an illusion, but it is not an illusion that is going to be whisked away by any amount of scientific explanation or heartfelt rhetoric. It is persistent. The magical self is here to stay.

I'm a case in point. I've devoted years of my life to the study of conscious will; I've written a book on how people experience an illusion of conscious will; I've taught seminars in the topic; and my lab has produced an array of experiments examining the wellsprings of the experience of will. If the illusion could be dispelled by explanation, I should be some kind of robot by now, a victim of my own nefarious schemes. No self, no magic, no inner agent. Yes it's true, when I'm on the dance floor I may look a bit robotic to some—but I'm happy to report that despite my personal flurry of illusion-busting, I remain every bit as susceptible to the experience of conscious will as the next person. It feels like I'm doing things.

Why is the illusion of conscious will so persistent? One possibility is that the processes of causal inference we use to establish our feelings of will are not entirely logical. That is, information about the operation of causal consistency, priority, and exclusivity may not combine in a way that is simple or additive. The persistence of the illusion of conscious will suggests that we can experience this illusion at times even when we are faced with causal information that indicates we are *not* the authors of our action. The fact that people in our various experiments could accept authorship for events over which they clearly had no control—such as another person's movements, or an event happening at a distance—suggests that there is a magnetic quality to the experience of will, an attractiveness that draws us toward it despite clear evidence that it cannot be true.

Perhaps the best evidence of the illusion of agency swamping causal logic appears in a set of experiments conducted by Woolfolk, Doris, and Darley (2006). These studies examined the perception of others rather than self-perception, but they make the point nonetheless. For these experiments, people were asked to judge the moral responsibility of a person who committed a killing—he shot a man. Some participants were told that the shooter did so under overwhelming duress; either he had been forced to shoot while he himself was standing in the aim of a firing squad with automatic weapons, or he had been forced to shoot because he had been given a drug that rendered him utterly at the mercy of others' commands. And as we would expect, these external forces were judged to lessen his moral responsibility for the shooting. In line with theories of discounting in causal attribution (Kelley, 1972; McClure, 1998), people saw these forces as reducing his moral culpability.

The intriguing finding of these studies was what happened when some people were told that he *meant* it. Some participants read that the shooter wanted to do it (because the victim had done harm to him in the past), whereas others were not so informed. The result: The shooter who wanted to do it was judged more responsible than the one who didn't—*even when he performed the act in the presence of overwhelming external causation (the firing squad or the drug) that clearly made him do it!* These findings suggest that people judge moral responsibility in a way that does not follow from a simple model of causal logic. It is as though we would hold the moon responsible for orbiting the earth if we could somehow discern that the moon wanted to do this—despite the clear operation of the laws of gravity and motion.

It may be that the illusion of conscious will is persistent because we honor so deeply what people mean to do that we readily overlook the causal forces that have impinged on them to force their action. In terms of the apparent mental causation theory, it may be that consistency information trumps exclusivity information. Merely having a thought or desire to perform an action is seen as the beginning of a magical connection from self to action regardless of what is going on in the world outside the self. The persistence of the illusion of conscious will could be due to processes of social evolution that have led us to a profound appreciation of what people think about what they do. We are drawn to the illusion of conscious will because we value so highly the cognitive previews of actions that our minds can provide for us.

The Wellsprings of Magic: Illusion Evolves

If conscious will is such a mesmerizing personal illusion that it persists under every sort of explanatory insult, perhaps that's because it has been bred into us by an

evolutionary process that has found the magic of the experience of will to be useful for the propagation of members of our species who fall under its spell. Perhaps people who don't experience conscious will fail to compete in society, fail to mate, or fail to parent successful offspring. Perhaps the societies they fashion work ineffectively and fail to survive. In contrast to Dennett's (2003; this volume) suggestion that freedom evolves, perhaps it is the illusion of freedom that evolves.

Why would an illusion evolve? The way to start thinking about this is to examine the downstream effects of the illusion. What personal or social consequences does the illusion of conscious will produce that might have proved sufficiently beneficial to individuals or to social systems that are advantageous to individuals that the whole mental apparatus for producing this illusion had to arise as an adaptation?

There are at least three such consequences of believing in the magic self, but deciding among them is difficult. As we well know, evolutionary arguments can be hard to test because they typically take the form of after-the-fact interpretations: How did a current trait of a living organism arise as an evolutionary adaptation to a particular environment by an ancestor organism that didn't have that trait? There are ways to test these stories (Cleland, 2001; Conway & Schaller, 2002), but the first step is telling the stories. So in this final section of the chapter, let's consider three possible avenues for the evolution of magic: *social signaling*, *social task allocation*, and *social control*.

Social signaling. As a first step in the functional analysis of conscious will, we need to pare it down to its basic element: The experience of willing an action is a conscious indication that one is the cause of the act. The experience is helpful as a marker of authorship, what I've called an "authorship emotion" (Wegner, 2002). Like

the anger that translates an event of losing something into an experience that accentuates the loss in one's mind, conscious will translates an event of doing something into an experience that accentuates one's likely causal role in mind. Regardless of whether the experience of will is technically correct or not, it highlights in one's own mind the events in the world that seem to have been authored by self. The further usefulness of the experience of conscious will is that it gives us something we can communicate to others—a feeling of doing that we can then use to tell the world what we believe we have done.

The ability to give these self-assessments does not come without cost. The experience of willing is an *addition* to the usual processes that create action—a lean-to built beside the main barn where the actions are made. This add-on is an authorship module of mind that visits experiences of authorship on some of our behaviors, and fails to produce this experience for others. Many of the things we do might well occur without experiences of authorship, and in fact, this is true for a wide range of behaviors commonly described as automatic (Wegner, 2005). They simply happen and we don't pay much attention to the fact that they issued from us. Indeed, we may not even notice who did it when an itch gets scratched or our position gets shifted in our chair. The authorship of such actions escapes our attention, and a range of such habits and rituals occur without the benefit of mental processes accompanying them that keep track of who did it. It is in the case of actions we feel will have freely willed that we have an experience that reminds us that we did them—and so tells us that these particular movements were not the result of external events, the movements of others' bodies, or the machinations of others' minds.

A key element needed for the feeling of will is the ability to think about our actions, ideally before they happen. This human capacity to experience mental *previews* of what we will do is, of course, the foundation of the common sense theory of intention—the idea that our thoughts truly cause our actions. But if we set aside this common sense theory in favor of the theory of apparent mental causation, why then would people need previews? If thoughts are not really needed prior to actions, why would evolution go to the extensive trouble to provide them for us? It could be that these thoughts arise not to cause action, but to signal its possible occurrence to us—a kind of warning that tells us what we can expect our bodies to do. Our thoughts about what we do may be part of a system for self-prediction. It would be impossible to tell a self-predicting system from an intending system if the self-predictions were accurate enough.

Self-predictions could be useful, much as it is useful to have dashboard gauges that tell us when the fuel is empty, or oven temperature lights to tell us when the turkey is baking. The ability to think and talk about our actions well in advance of their occurrence is of particular utility for social purposes: We can tell when someone else might do something bad or good for us because their minds have handy self-prediction functions that have prompted them to tell us what they intend. When that angry fellow at the bar says he's going to break a pool cue over your head, you have a signal indicating what might happen. The interesting feature of such signals is that they can save *both* of you the trouble of actual physical harm. Communications of intention serve the purpose of making many potentially costly social actions unnecessary because the statement itself causes preventive responses.

Darwin (1872) pointed out that displays of dominance and submission are very useful to animals because they regularly take the place of actual conflict. In humans, as in dogs, exhibitions of aggressive intention can trigger us to roll over on our back and expose our tummies in a gesture of good will and desire for scratching—and thus prevent the mayhem and mutual danger that might have ensued. Perhaps previews of our actions come to mind so we can convey our likely behavior to others before it happens, and so signal our way out of social emergencies before they occur. Conscious will and intention may be much like turn signals on automobiles, features that have arisen to prevent accidents and so save the car for another trip tomorrow.

Social task allocation. What else does conscious will do? Another evolutionary story could be told about the role that our self-knowledge of action tendencies can play in helping us to choose tasks that will be useful to us, and that will also be useful to society and so bring us social rewards such as paychecks and opportunities to mate. Finding the right niche for our own special talents may be facilitated by experiences of conscious will.

In many social animals, particularly social insects, the job a given organism will do in the society is determined by its inherited morphology. The phenotype of a given caste of ant, for example, may be that it has an unusually large head. This makes ants of this form very useful to the whole colony as specialized soldiers whose job is to act as doors, so ants of this type spend all their time around the entrances, using their heads to shut the passageway to intruders and opening the way only when ants of their colony prompt them by tapping antennae on their big noggins (Holldobler & Wilson, 1990). We humans have not evolved such a system for task allocation in our species, so even those

with very large heads seldom get work as doors. Instead, we develop systems for task allocation determined in part by self-assessed expertise. People who discern that they have certain abilities often take roles in society that take advantage of those abilities. If a person has regularly experienced a sense of willing associated with particular actions—hitting a baseball, for example, or arguing a point—that person may well become a ball player or an attorney as a result of the self-knowledge that this experience has provided. Conscious will conveys the sense that "I can do this" and human social organization has use for people who can select tasks they can do, even if that work involves no more than blocking doors.

Much of the psychological literature on perceived control has focused on this idea—that it is good to perceive control over those things one does indeed control, and not good to perceive control over those things that are in fact out of hand (Baumeister, Heatherton, & Tice, 1995; Burger, 1989; Folkman, 1984; Haidt & Rodin, 1999; Peterson, Maier, & Seligman, 1993). The experience of conscious will provides an anchor of sorts, an internal point of reference that is the mind's "best estimate" of whether the event in question might indeed be traceable to oneself. Although this estimate can never be correct scientifically in all respects (Davies, in press), it certainly has its uses. The magic self is a natural guide to the roles one can play most effectively in life and in society. It tells us what we can and cannot do.

Social control. Another social function for conscious will is to ready individuals to accept responsibility. As the theory of apparent mental causation suggests, the feeling of willing may be a poor indication of true causal responsibility, as this feeling can come and go in error. However, the theory also suggests that the feeling of willing that does

arise in an individual for any action will compel that individual to accept personal responsibility for that action. Right or wrong, such responsibility acceptance then prepares the person for the experience of moral emotions such as pride for right action and guilt for wrong action. The experience of conscious will provides a unique inner signal, a first-person experience of responsibility that makes the person "own" morally-relevant actions.

There are many instances when people don't accept complicity in actions that have moral overtones. Sometimes people are deceptive about these things—they deny performing crimes or they claim good deeds as their own—but the responsibility acceptance fostered by apparent mental causation guarantees that they will not be universally deceptive. The feeling of conscious will resonates in the person's mind even when there are clear reasons to lie about authorship of an action, and the feeling may guide the person to admit to wrongdoing (or own up to the lack of right-doing) even in the face of these reasons. This inner feeling of doing can behave like a "conscience," weighing in to make the person truthful about moral actions. Even if that truth is the expression of an illusion, it is an illusion that derives from the person's own best guess about the authorship of the action.

The acceptance of personal responsibility is a useful step toward successful social control of individual moral action. When society delivers its third-person judgments of responsibility—as when the law says someone is a criminal, or when a parent praises a child for helping with dinner—the individual's first-person feelings of responsibility will incline the individual to comply with these external judgments of culpability. Being sent to jail, or even being given a humanitarian award, would be difficult if we didn't

authentically feel that we were the ones who had authored the moral action that had earned us those desserts. Quite simply, our inner feelings of doing give some license to the social world to hold us responsible.

Commentators who worry about the fate of legal and moral responsibility in a world that recognizes the illusory nature of conscious will have not come to appreciate the profound impact of the person's own moral sense. It matters far less for moral purposes what a person really did than what that person feels responsible for doing, and it is the feeling of responsibility that thus must be cultivated by social evolution. Our authorship processing modules may be good enough, often enough, that we typically get our responsibility roughly right. Like bonobos who hold each other responsible for food theft, or who know who should be punished for free riding (Boehm, 2001), humans who have a ready sense of their own complicity in right and wrong actions are likely to work effectively in social settings and survive some trials of social evolution.

The acceptance of individual responsibility for moral actions is essential for the exercise of social control. Indeed, the actions for which people experience free will are typically those actions that are most likely to be susceptible to modification by social consequences. In the study of animals, voluntariness is usually defined in terms of behavior modifiability (Passingham, 1993). Likewise in humans, actions we perceive as voluntary are also actions that are susceptible to modification through reinforcement. The things people feel apparent mental causation for, then, are those that are likely to be the focus of attempts at social control. If you say you can postpone a sneeze, people may pay you to do it when they want quiet—whereas if you find the sneeze inevitable, few attempts at control will come your way. This is nothing to sneeze at, however, as the

identification of behavioral candidates for social control is serious business when it comes to criminal or immoral behavior. Far from eliminating responsibility for our behavior, then, the mental processes that produce the illusion of conscious will seem to be part of the mechanism that creates such responsibility and makes behavior more open to modification.

Conclusion: Living with an Illusion

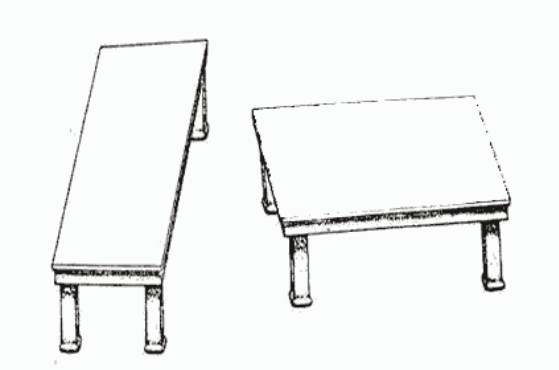
The life of the magician is not easy. Like Harry Potter, each of us must make sense of our amazing tricks of action and somehow fit our understanding into a sensible view of the world. This elliptical-peg-into-trapezoidal-hole problem is not an easy fit, and anguish surrounding the issue of free will and determinism echoes throughout philosophy and psychology. The chorus of discordant voices in this volume reveals that psychology continues to struggle with this conundrum. The uneasy solution suggested in this chapter involves learning to live with the magic.

The solution begins with recognizing that the magic is a little show we put on for ourselves. The sense of what we consciously will is only part of an authorship estimation system of mind, which can thus be mistaken—and a number of experiments in my lab and others suggest that the experience of will can not only be off, it can be wildly wrong. The second step toward solving the problem of free will is recognizing that, even when it is wrong, the magic we perceive in ourselves trumps other explanations *in our own minds*. Like that visual illusion of disparate tables that just won't quit even when we know it is an illusion, the magic self rules our intuitions and won't be undermined by analyses of its workings. (This realization is heartening for those worried about what would be left if the magic were gone: The magic is here to stay.) The third step in

solving the problem of free will is discerning what functions this experience served in our biological and social evolution—and which of these might be so crucial that we absolutely had to have this magic installed in our heads. The third step is just beginning, as the understanding of the evolution of mind will take time. Until we take that step, we should be content to continue conducting scientific psychology to understand the mind, secure in the realization that the scientific discovery of our inner processes will never make us any less magical to ourselves.

Figure 1.

The Shepard Tables



References

Aarts, H., Custers, R., & Wegner, D. M. (2004). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and Cognition, 14*, 439-458.

Aarts, H., Wegner, D. M., & Dijksterhuis, A. (2006). On the feeling of doing things: Dysphoria and the implicit modulation of authorship ascription. *Behaviour Research & Therapy, 44*, 1621-1627.

Angel, L. (1989). *How to build a conscious machine*. Boulder, CO: Westview Press.

Bargh, J. A. (2004). Bypassing the will: Towards demystifying the nonconscious control of social behavior. In R. Hassin, J. S. Uleman & J. A. Bargh (Eds.), *The new unconscious*. New York, NY: Oxford University Press.

Bargh, J. A. (this volume). The primacy of the unconscious. In J. Baer, J. C. Kaufman & R. F. Baumeister (Eds.), *Psychology and free will*. New York: Oxford University Press.

Bargh, J. A., & Barndollar, K. (1995). Automaticity in action: The unconscious as repository of chronic goals and motives. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The Psychology of Action* (pp. 457-481). New York: Guilford.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trotschel, R. (2001). The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology, 81*, 1014-1027.

Baumeister, R. F. (1997). *Evil: Inside human cruelty and violence*. New York: W. H. Freeman.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1995). *Losing control.* San Diego, CA Academic Press.

Blackmore, S. J., Galaud, K., & Walker, C. (1994). Psychic experiences as illusions of causality. In E. Cook & D. Delanoy (Eds.), *Research in parapsychology 1991* (pp. 89-93). Metuchen, NJ: Scarecrow.

Boehm, C. (2001). *Hierarchy in the forest: The evolution of egalitarian behavior*. Cambridge, MA: Harvard University Press.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Brown, J. W. (1989). The nature of voluntary action. *Brain and Cognition, 10,* 105-120.

Burger, J. M. (1989). Negative reactions to increases in perceived personal control. *Journal of Personality and Social Psychology, 56,* 246-256.

Cannon, W. B. (1942). "Voodoo" Death. *American Anthropologist, 44,* 182-190.

Claxton, G. (1999). Whodunnit? Unpicking the "seems"of free will. *Journal of Consciousness Studies, 6,* 99-113.

Cleland, C. E. (2001). Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science, 69,* 474-496.

Conway, L. G., III, & Schaller, M. (2002). On the verifiability of evolutionary psychological theories: An analysis of the psychology of scientific persuasion. *Personality and Social Psychology Review, 6,* 152-166.

Darwin, C. (1872). *The expression of emotions in man and animals*. New York: D. Appleton and Company.

Davies, P. S. (in press). What kind of agent are we? A naturalistic framework for the study of human agency. In D. Spurrett, D. Ross, J. Kincaid & L. Stephens (Eds.), *Distributed cognition and the will*. Cambridge, MA: MIT Press.

Dawkins, R. (1998). *Unweaving the rainbow*. New York: Houghton Mifflin.

Dennett, D. (1984). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Dennett, D. (2003). *Freedom evolves*. New York: Viking Press.

Dennett, D. C. (1994). The practical requirements for making a conscious robot. *Philosophical Transactions of the Royal Society, A349*, 133-146.

Dennett, D. C. (this volume). Some observations on the psychology of thinking about free will. In J. Baer, J. C. Kaufman & R. F. Baumeister (Eds.), *Psychology and free will*. New York: Oxford University Press.

Denno, D. W. (2002). Crime and consciousness: Science and involuntary acts. *Minnesota Law Review, 87*, 269-400.

Folkman, S. (1984). Personal control and stress and coping processes: A theoretical analysis. *Journal of Personality and Social Psychology, 46*, 839-852.

Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions: Biological Sciences, 359*(1451), 1775-1785.

Haidt, J., & Rodin, J. (1999). Control and efficacy as interdisciplinary bridges. *Review of General Psychology, 3*, 317-337.

Holland, O. (Ed.). (2003). *Machine consciousness*. Exeter, UK: Imprint Academic.

Holldobler, B., & Wilson, E. O. (1990). *Ants*. Cambridge, MA: Harvard University Press.

Hume, D. (1739/1888). *A treatise of human nature*. London: Oxford University Press.

Jones, L. A. (1988). Motor illusions: What do they reveal about proprioception? *Psychological Bulletin, 103*, 72-86.

Kelley, H. H. (1972). Causal schemata and the attribution process. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 151-174). Morristown, NJ: General Learning Press.

Kelley, H. H. (1980). Magic tricks: The management of causal attributions. In D. Gurlitz (Ed.), *Perspectives on attribution research and theory: The Bielefeld Symposium* (pp. 19-35). Cambridge, MA: Ballinger.

Lynn, S. J., Rhue, J. W., & Weekes, J. R. (1990). Hypnotic involuntariness: A social cognitive analysis. *Psychological Review, 97*, 169-184.

Lynn, S. J., Weekes, J. R., Matyi, C. L., & Neufeld, V. (1988). Direct versus indirect suggestions, archaic involvement, and hypnotic experience. *Journal of Abnormal Psychology, 97*, 296-301.

McClure, J. (1998). Discounting causes of behavior: Are two reasons better than one? *Journal of Personality & Social Psychology, 74*, 7-20.

Michotte, A. (1963). *The perception of causality* (T. R. Miles & E. Miles, Trans.). New York: Basic Books.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67*, 371-378.

Milgram, S. (1974). *Obedience to authority*. New York: Harper & Row.

Miller, A. G., Gordon, A. K., & Buddie, A. M. (1999). Accounting for evil and cruelty: Is to explain to condone? *Personality & Social Psychology Review, 3*, 254-268.

Monahan, J., Hoge, S. K., Lidz, C., Roth, L. H., Bennett, N., Gardner, W., et al. (1995). Coercion and commitment: Understanding involuntary mental hospital admission. *International Journal of Law & Psychiatry, 18*, 249-263.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Passingham, R. E. (1993). *The frontal lobes and voluntary action*. Oxford, England UK: Oxford University Press.

Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science, 8*(3), 162-166.

Peterson, C., Maier, S. F., & Seligman, M. E. P. (1993). *Learned helplessness: A theory for the age of personal control*. New York, NY, USA: Oxford University Press.

Preston, J., & Epley, N. (2005). Explanations versus applications: The explanatory power of valuable beliefs. *Psychological Science, 16*, 826-832.

Preston, J., & Wegner, D. M. (in press). The eureka error: Inadvertent plagiarism by misattribution of effort. *Journal of Personality and Social Psychology*.

Pronin, E., Wegner, D. M., McCarthy, K., & Rodriguez, S. (2006). Everyday magical powers: The role of apparent mental causation in the overestimation of personal influence. *Journal of Personality and Social Psychology, 91*, 218-231.

Scassellati, B. (2002). Theory of mind for a humanoid robot. *Autonomous Robots, 12*, 13-24.

Seligman, C., Fazio, R. H., & Zanna, M. (1980). Effects of salience of extrinsic rewards on liking and loving. *Journal of Personality and Social Psychology, 38*, 453-460.

Shariff, A. F., Schooler, J. W., & Vohs, K. D. (this volume). The hazards of claiming to have solved the hard problem of free will. In J. Baer, J. C. Kaufman & R. F. Baumeister (Eds.), *Psychology and free will*. New York: Oxford University Press.

Shepard, R. N. (1990). *Mindsights*. New York: W. H. Freeman.

Sparrow, B., & Wegner, D. M. (2006). *The experience of authorship in coaction.*Unpublished manuscript, Cambridge, MA.

Spence, S. A. (1996). Free will in the light of neuropsychiatry. *Philosophy, Psychiatry, & Psychology, 3*, 75-90.

Thompson, S. C., Armstrong, W., & Thomas, C. (1998). Illusions of control, underestimations, and accuracy: A control heuristic explanation. *Psychological Bulletin, 123*, 143-161.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D. M. (2003). The mind's best trick: How we experience conscious will. *Trends in Cognitive Sciences, 7*, 65-69.

Wegner, D. M. (2004). Precis of The Illusion of Conscious Will. *Behavioral & Brain Sciences, 27*, 649-692.

Wegner, D. M. (2005). Who is the controller of controlled processes? In R. Hassin, J. Uleman & J. Bargh, A. (Eds.), *The new unconscious* (2 ed., pp. 19-36). New York: Oxford University Press.

Wegner, D. M., & Erksine, J. (2003). Voluntary involuntariness: Thought suppression and the regulation of the experience of will. *Consciousness & Cognition, 12,* 684-694.

Wegner, D. M., & Sparrow, B. (2004). Authorship processing. In M. Gazzaniga (Ed.), *The New Cognitive Neurosciences* (3rd ed., pp. 1201-1209). Cambridge, MA: MIT Press.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology*.

Wegner, D. M., & Wheatley, T. P. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist, 54*, 480-492.

Wilson, T. D., Gilbert, D. T., & Centerbar, D. B. (2003). Making sense: The causes of emoional evanescence. In I. Brocas & J. Carrillo (Eds.), *The psychology of economic decisions: Vol. 1. Rationality and well-being*. New York: Oxford University Press.

Woolfolk, R. L., Doris, J. M., & Darley, J. M. (2006). Identification, situational constraint, and social cognition: Studies in the attribution of moral responsibility. *Cognition, 100*, 283-301.