5

Ideal Agency: The Perception of Self as an Origin of Action

JESSE PRESTON AND DANIEL M. WEGNER

It is chiefly my will which leads me to discern that I bear a certain image and similitude of Deity.

—René Descartes, 1641/1901 "Meditations IV"

Each of us has ideals, things we would like to be. We would like to be happy, pretty, smart, brave, kind, or honest. Such desired characteristics are often appreciated in psychology as facets of a person's ideal self (Cooley, 1902; Rogers, 1961), and studies of the ideal self commonly focus on the particular selection of such ideals that an individual embraces. These ideals are all descriptions of the "good person." In this sense, they all skip right over an important point. There is a far more fundamental ideal that often goes unnoticed, a basic prerequisite for even being a person at all. This is the ideal of agency. Quite simply, we would rather be agents than objects.

Yes, of course, anyone would rather do things than be a block of wood. It turns out, though, that the ideal of agency is far more subtle, pervasive, and even insidious than a mere desire to be active. This chapter is about how the desire to be an agent influences human psychology. As suggested by Wegner (2002), aspiration to the ideal of agency is a profound force that underlies and connects many patterns of human thought and behavior. Ideal agency is involved in how people come to believe they consciously will their actions, it operates in the processes (usually understood in terms of cognitive dissonance) that underlie the perception and confabulation of intentions, and it functions as well to produce the misperception of actions as successful. It takes a lot of work not to be a block of wood, and the aspiration to the ideal of agency prompts an unsettling array of psychological distortions in service of this goal.

To examine the influences of ideal agency, we will consider first what it means to be an agent of any sort. There are basic properties of agency that can inform the analysis of how humans aspire to this goal. This analysis allows us to understand how an agent could be more or less ideal. With a grasp on the dimension of ideal agency, then, we will examine each of three essential paths toward the ideal—the confabulation of intention, the experience of an illusion of conscious will, and the misperception of action.

AN ANATOMY OF AGENCY

Agents are entities that cause events through self-movement. By this definition, humans, dolphins, gerbils, grasshoppers, robots, bean sprouts, certain artificial intelligence (AI) programs, and God Almighty all hold in common the property of agency. Agents can be contrasted with objects—entities that do not cause events. Objects include the billiard ball, as well as the aforementioned block of wood, and many other items in the universe that appear to move only when moved upon. In a sense, the unique property of agents is the inscrutability of their motion. Agents are first causes or uncaused causes. Authorship is ascribed to agents as the origin of action because their prior causal history is unclear.

The notion of agency is so central to human psychology that it is regularly assumed rather than examined. We readily perceive agents all around us and seldom mistake them for objects. The philosophy of mind reveals agency to be a major human preoccupation. Much thought and theory is aimed at understanding agent causation and how it can be understood as distinct from object causation (e.g., Anscombe, 1957; Brand, 1984; Davidson, 1980). And psychologists have long believed that the perception of persons is dominated by the theme of agency rather than by a theme of mechanism or simple physical causation (DeCharms, 1968; Heider, 1958; Heider & Simmel, 1944).

The fundamental character of agency perception is suggested by the finding that judgments of agency appear to be snap decisions, based on visual cues such as biological motion (Blakemore & Decety, 2001) and agentlike appearance, such as eye spots (Baron-Cohen, 1994). The perception of agents is deeply ingrained in our minds, a property of our early perceptual systems that is amplified and refined as we grow to develop mature "theories of mind" (Premack & Woodruff, 1978; Wellman, 1992). We not only perceive the real agents around us but we also make up a spirit world of angels, devils, gods, and other supernatural agents to account for events that we cannot otherwise explain (Barrett, 2000; Boyer, 2001). Indeed, Guthrie (1980) suggested that we have evolved to perceive agency because it is of crucial importance to detect other agents in order to survive. After all, it is much better to mistake a boulder for a bear than to mistake a bear for a boulder.

When we perceive agency in ourselves, others, or imagined others, what is it that we think we are seeing? Are there common features of things that cause themselves to behave? What, then, makes up an agent? In a discussion of the characteristics of artificially intelligent agents (computer programs that "do

things"), Russell and Norvig (2003) provided a useful three-part architecture of agency that applies broadly to both natural and artificial agents. In essence, agents can be described in terms of the operation of three key features: sensors, processors, and actuators (Russell & Norvig, 2003). Agents input information through sensors, operate on the information through processors, and output behavior through actuators.

Sensors are the means by which the agent knows about its environment. Just as people sense and come to know their environment through sight, smell, hearing, taste, and touch, an AI agent such as a robot could sense its environment by detecting variation in heat and light, for example. The *processor* is the mechanism by which the agent chooses a course of action based on the input. Russell and Norvig discussed three different kinds of processors with which an AI agent could be programmed: if–then rules (e.g., if the car in front of you is braking, then initiate braking), goal-based processors (which are based on trying to get to a set goal), and utility-based processors (which select actions that are "good" for the agent—have high utility—without having a specific goal). The human mind might potentially implement any of these sorts of processors (Carver & Scheier, 1998; Newell, 1973; Powers, 1990), whereas software agents are typically programmed with just one, as appropriate to their specific task. Regardless of the specific programming, however, the AI processor, like the human, uses the sensed input to select some actions over others, and that selection is based on criteria applied by the agent. The *actuator*, the third property of the agent, is the means by which the agent acts upon the environment. People have the limbs, nerves, and muscles of their bodies to get what they want. An AI agent could have mechanical limbs or other tools at its command that provide the same function. In the simplest AI agent, the actuator could merely be some computational function that changes stored information in memory or turns on a pixel on the computer screen.

Despite the proficiency of computer programs to calculate and even outplay chess masters, AI researchers have yet to actually build an AI agent that approaches the sophistication of human agents. To be sure, there is a hierarchy among agents, in that some agents are better than others. The hierarchy is not a matter of who is flesh and who is metal. The hierarchy of agents involves agents that are better or worse at the three defining features of agency: sensing, processing, and actuating. The best possible agent would be the one that performs all three functions most effectively.

IDEAL AGENCY

The ideal agent thinks about the situation and action, chooses what will enhance its happiness (or achieve its goal), and then successfully performs the action that is needed. The ideal agent, in other words, senses ideally, processes ideally, and actuates ideally.

There are agents all around us that fall short of this ideal. There is the copy machine that can nicely sense the presence of an original to be copied, can process

the input when we press the button for five copies, but that then actuates the copies into a serious paper jam. There is the dollar-bill changer that can't sense our rumpled bill and keeps spitting it back. There is the Web site that accepts our order and successfully sends us the item we have purchased but also processes the order repeatedly and sends us the item again and again. Modern life is filled with simple agents gone wrong, and this is before we even get to the errors of flesh-and-blood agents. There are dogs who won't fetch; parrots who only repeat vile language; and all the bad drivers, incompetent waiters, and insolent store clerks who populate our daily lives. Their breakdowns of agency can be traced alternately to poor sensing, poor processing, or poor actuating.

In contrast to all these bad agents, there is the perfect agent. In some cultures, this ideal is human, but in many it is God. The dominant Western concept of God portrays an agent that is omniscient, benevolent, and omnipotent. To be omniscient means that God knows all and is aware of everything. God's sensors are therefore perfect. To be benevolent means that God wills good things to happen. God's will is therefore perfect in that God knows exactly what the right thing is that should be done in order to cause events that achieve the best possible goals. In this sense, God's processor has perfect utility. And to be omnipotent means that God can do anything—can affect anything and change it in any way that God likes. God's actuators are therefore perfect, in that the outcome desired is exactly the outcome that God achieves. Deities with lesser talents appear in religions around the world, but their godliness is typically accompanied by a heightened level of one of these characteristics—lesser gods can sometimes at least sense more acutely, will more benevolently, or act more powerfully than mere mortals. Thus, the concept of God is an agent more ideal than the human, and sometimes perfectly ideal.

We humans fall somewhere between faulty vending machines and the God of infinite perfection. More often than we would like to admit, we behave without full knowledge of our situation and how it pertains to what we are doing. Thoughts about our situation and what should cause us to act would require perfect sensors, and we don't always have them. We also may have knowledge and perform an action but fail to choose the proper course. Willing our actions correctly would require perfect processors, and we don't always have them. Finally, although we may have the thought and the will, our capacities to act may fail us. The ability to perform the actions we envision would require perfect actuators and, sadly, we don't always have these, either. However, we know what ideal agency should be, and we are motivated by our conception of this ideal to assimilate our actions to it. Thus, we fill in the blanks: When possible, we assume that we have thought, will, and action that exemplify ideal agency.

IDEALIZATION PROCESSES

The fact that our humble bodies and minds do not always follow the ideal thought–will–action model means that we often find ourselves with only some fragments of the ideal. However, these failures of agency could have the effect of

making the agent more likely to fail in the future. Experiences of loss of control and expectations of failure can lead to learned helplessness (Abramson, Seligman, & Teasdale, 1978; Dweck & Leggett, 1988) and self-handicapping behavior (Arkin & Baumgardner, 1985; Berglas & Jones, 1978), turning judgments of the self as a faulty agent into a self-fulfilling prophecy. It may be better for the agent to not pay too much attention to the "blips" in the system and believe that everything is operating as it should. In aspiring to the ideal of agency, then, we infer that the missing components of ideal agency are actually in place. People do more than merely deduce that the aspects of ideal agency are present; they also actually perceive them to be there. If we cannot directly observe any one of the components, we may infer that it has functioned and fill it in with the relevant information, just as the blind spot is filled with the appropriate background by the visual system.

Certainly there are times when we are content to be objects. Agency is not even an issue for whole classes of behavior (e.g., knee reflexes, eye saccades, sneezes, autonomic nervous system responses), and it also is not a concern when our bodies are moved about by external forces (e.g., lifted by an elevator, bumped by a door). And for events that occur far from our bodies, we may have an initial filtering system that screens out those for which we could not be agents at all. However, for a range of actions of our bodies that we commonly understand as voluntary (Passingham, 1993), possibilities of agency are always present. For such actions, we attempt to apply a template of ideal agency, and when it seems to fit, we assume that thought, will, and action are all in place.

This means that whenever any of the components of ideal agency is absent, we tend to feel a certain tension and incompleteness. Something seems amiss. For example, when you walk into the kitchen and realize on arrival that you don't remember what you intended to do when you got there ("Was I coming in here for cookies? Milk? A double whiskey with beer chaser?"), something seems wrong. You performed an action (walking into the room), it somehow felt willed (you know you weren't pushed, anyway), and yet you don't have the thoughts that would tell you what you had sensed to lead you in this direction. You don't know what you were doing. The case of less-than-ideal agency feels somehow distressing. This observation can be generalized to a wide range of action: We usually assume that all three components of ideal agency will be present, and we find it discomfiting when any one is absent.

Beyond a basic feeling that something is wrong, the ideal of agency prompts people to make motivated inferences. Specifically, whenever some component of ideal agency is missing, the person is motivated to infer that the component is indeed present. This inference engine is weakest when all three components of ideal agency are missing. Thus, when a person has no relevant thought, will, or action, the person is quite unlikely to infer that any or all of these are present. Only rarely do people assume they were the agents of actions performed by others, which they did not conceive of in advance, and which they did not feel that they willed. Admittedly, some alien-control hallucinations in schizophrenia may have these properties (Stephens & Graham, 2000), but they are rare. Claiming authorship of environmental events at a distance is not the normal human proclivity and

only occurs when there are serious breakdowns in the brain and mind processes responsible for authorship processing (Frith, Blakemore, & Wolpert, 2000; Wegner & Sparrow, in press).

When one component of agency is present, somewhat stronger inferences of ideal agency are made. A person who thinks about an action (say, telling off the boss) may have a tendency to perceive that the action happened (noticing, e.g., that she actually had said something the boss might not have liked) and may also feel that she willed this action (even to the point of feeling guilty about her imagined slur)—when these idealized accompaniments of the thought actually did not occur. By the same token, perceiving that an action has occurred (such as having hurt someone) might lead the person to infer that this action was thought about in advance and even that it was willed. And feelings of will ("I want to do better in school") might bring about the inference that relevant thoughts had crossed one's mind ("I studied for the test") and even lead to a tendency to perceive that the action had happened (as in the case of a C grade misremembered as an A).

These motivated inferences become far more likely, and far more powerful, when multiple components of ideal agency are already in place and just one is missing. Thought, will, and action are presumed to co-occur so profoundly that when we get two out of three, we assume we've won the whole match. When will and action are present, we infer thought; when thought and action are present, we infer will; and, when thought and will are present, we infer action. Each of these three patterns of motivated inference leads to a different range of distortions in self-perception, some of which are already widely known and studied in psychology and others of which are only beginning to be discerned. Different circumstances may lead to different inferences' being made. However, whatever is most ambiguous and unclear will typically be altered to match the more unambiguous components of the ideal.

The three strong inferences from ideal agency are the focus of the remainder of the chapter. As shown in Table 5.1, these inferences each follow when evidence of ideal agency is present for two components of the ideal and is not present for the third. They include *intention confabulation* (inference of premeditative thought), *apparent mental causation* (inference of the will), and *action misperception* (inference of action).

Intention confabulation occurs when we can observe that we have performed an act and feel as though it was willed but cannot access the thoughts or intentions

TABLE 5.1. Strong Inferences of Ideal Agency

	Component of ideal agency		
Type of inference	Thought	Will	Action
Intention confabulation	Inferred	Observed	Observed
Apparent mental causation	Observed	Inferred	Observed
Action misperception	Observed	Observed	Inferred

that led us to perform the act. An intention is then inferred after the fact that suitably explains why we would have wanted to perform the action in the first place. Apparent mental causation occurs when we observe that we have performed an act and had intention to perform the act but do not have access to willing that act. Knowing that we had prior thoughts and sensations that suggested the intention, and seeing that the action was indeed performed, leads to the feeling that the act arose out of our will. Action misperception occurs when we have intention to perform an act and feel as though we have willed the act but do not have access to the act itself. Whether or not the action occurred as intended may be altered in our minds so that we view the outcome to be more successful. In what follows, we discuss each of these inferences in turn and review the evidence indicating some of the conditions under which each of the inferences has been observed.

INFERENCE OF PREMEDITATION: INTENTION CONFABULATION

In a scene in the 1985 film *Pee-Wee's Big Adventure* (Shapiro, Abramson, & Burton, 1985), the title hero is seen riding his beloved bicycle through the streets, doing a number of daring tricks. At one point, a trick goes awry, and Pee-Wee is catapulted off his bike. He flies hundreds of feet through the air, landing in front of a group of neighborhood kids. Dusting himself off, our hero dryly comments "I meant to do that," and walks off.

We all fly through the air in our own special ways only to land on the ground and claim that we meant to do that. A wealth of research has caught people making this particular inference of ideal agency in a variety of situations. But we should not be too quick to chastise ourselves for such selfish delusions; we also have an astounding lack of insight into the reasons for all sorts of actions—good, bad, or neutral. What is particularly interesting is that despite the fact that the reasons for what we do are often hidden from ourselves, the tendency to report reasons for what we do remains unimpaired (Wilson, Hodges, & LaFleur, 1995; Wilson & Schooler, 1991).

It appears as though we are unable to admit or understand that we may have acted without knowing why. Apparently, it is unpleasant to think that the things we do might not emerge from our personal resolutions to do them. Making ourselves do things is how we get all the things in the world we cherish, how we interact with the people with whom we have our most meaningful relationships, and how we do the jobs and the hobbies that define our identities. All these activities should begin from our desires and intentions to do them; otherwise, what could all of these behaviors be for? Thus, if we have performed some behavior, it surely must be for a reason, even if we only make that reason explicit after the act. The motivated inference of reasons and thoughts relevant to action has been termed *intention confabulation* (Wegner, 2002), and it occurs whenever we generate reasons to explain our actions post hoc and believe that those reasons were our intentions all along.

Confabulation by Children

There is evidence that young children are particularly prone to confabulate memory for intentions. Schult (1997) observed this in a study in which children tossed a beanbag toward three colored buckets. They were asked to name in advance which color bucket they wanted to hit and were given a color chip to remind them of their choice. Hidden at the bottom of one bucket was a prize—a picture signaling that they had "won" (and that they could mark on a score sheet). After each time they hit a bucket, the children were asked, "Which one were you trying to hit?" The interesting case here, of course, is when the child "wins" by hitting a bucket other than the color they had said they wanted. And, indeed, when the intended bucket was missed but a picture was found, the 3-year-olds answered incorrectly more often than the 4- and 5-year-olds. Like Pee-Wee, the youngest children often claimed they had been "trying to hit" the winner all along. Similar results have been reported by Abbott and Flavell (1996).

Cognitive Dissonance and Self-Perception

A wealth of evidence demonstrating intention confabulation comes from research on the theories of cognitive dissonance and self-perception. The phenomenon at the heart of these theories is that people report their attitudes to be consistent with the actions they perform, even if their prior attitude was in fact inconsistent with the action (Bem, 1972; Bem & McConnell, 1970; Cohen, 1962; Festinger, 1957; Wicklund & Brehm, 1976). The theories differ in the mechanism they use to explain this phenomenon. According to cognitive dissonance theory, the change in attitude results from the conflict a person feels when he or she holds two inconsistent cognitions. Festinger (1957) argued that people believe that their actions ought to arise from their attitudes. If their actions and attitudes are incongruent, this causes tension. As a result, one of the two cognitions is altered to match the other. In general, it is easier to deny an attitude than an action, so people change their attitude to match the behavior. According to self-perception theory (Bem, 1967, 1972) there need not be a prior attitude at all. Bem suggested that we often are in the same position as an observer of actions; we don't know how we feel until we see ourselves acting in some way. We believe that our attitudes and actions should be consistent, but sometimes it is only our actions that inform us of our attitudes. The conflict between these theories is not critical for our purposes, and in fact it has been quite nicely resolved (Cooper & Fazio, 1984). What matters is the common point—the idea that people will confabulate prior intentions based on their actions. This confabulation relies on the feeling that the action was freely willed by the actor, not induced by others.

In one of the first experiments on cognitive dissonance, Festinger and Carlsmith (1959) had research participants engage in an extremely dull task—turning pegs on a board. The participants were then asked to tell the next participant that this dull activity was fun and interesting. For their PR work, participants were paid either $1 or $20. Following this, participants were then asked to report their true feelings about the joy of peg turning. The people who were paid $20 reported

that the activity was actually quite dull and boring, despite what they had told the confederate. But apparently being poorly paid has its benefits, because the people who were paid $1 reported that turning pegs was enjoyable. Being paid $1 to do anything is not a very good justification for anything, certainly not for telling a lie to a stranger who has never caused you any harm; $20, on the other hand, is much more motivating, and one could justify telling a lie in order to get a healthy amount of cash in hand. What is interesting is that in no case did anyone refuse to lie for the experimenter. Compliance is very easily induced in participants, and people are remarkably eager to go along with the polite requests of a scientist. It appears, then, that in all cases participants confabulated an inaccurate intention for their action. In the case of participants paid $1, they believed that their endorsement of the activity stemmed from their true enjoyment of the activity. In the case of participants paid $20, they believed their endorsement stemmed from the generous monetary compensation. However, in all cases, their endorsement was simply compliance with the experimenter, possibly to be helpful to the experimenter or because they were submitting to authority. However, none of the participants were aware of the powerful forces of social pressure, and so they confabulated an intention to explain their behavior.

Induced compliance turned out to be very useful in this line of research, for it made the participants feel that they were acting of their own free will. In another study, Linder, Cooper, and Jones (1967) had students write an essay that contradicted their held attitude (to ban speakers at their college). As in the Festinger and Carlsmith (1959) study, the payment for this task varied. For some participants, the pay for writing the essay was announced as 50¢, whereas for others it was $2.50—a tidy sum back then. A further variation in the experiment manipulated the perception of choice: Some participants were led to believe that they had considerable personal choice with respect to whether to write an essay—the experimenter explained at the outset that after he had described the study they could decide for themselves whether or not to write the essay. For others, there was no such emphasis placed on choice.

As expected, everyone complied with the request. The effect of payment was also replicated: Those who wrote the essay that ran against their own opinion for 50¢ endorsed it more afterward than did those who wrote for the larger sum of $2.50. However, compared to participants who were not given the sense of choice (an illusory sense, because everyone did write the essay), those who perceived choice showed a classic cognitive dissonance effect. Performing a dissonant act under conditions designed to arouse a feeling of choice made people become more positive toward the topic they were paid less to espouse. When the coercion of the experimenter was overt and the experience of willing the action was thus undermined, the participants did not need to confabulate an intention to write the essay. Their action was accounted for by the experimenter's will. However, when participants were given an illusory sense of choice, they came to feel as though they were in favor of the attitude they presented in the essay. Why else would they "choose" to write such an essay?

These studies and many others demonstrate that people change or create their attitude to match the behavior they have performed. But does this necessarily

mean that they have confabulated a prior intention? It may very well be that behaving in a way has made them see the error of their ways and persuaded them to change their views. However, the researchers in the early dissonance studies reported that not only do people report a new inconsistent attitude but they also fail to recall that their prior attitude was different from their present one. Bem and McConnell (1970) addressed the issue of prior-attitude awareness. They asked people who had written counterattitudinal essays to report not their final (post-essay) opinion on the issue but their prior opinion. People couldn't do this. Instead, these reports of prior attitudes mirrored faithfully the standard dissonance effect: Participants led to believe that they had a high degree of choice in whether to write the essay reported that they had agreed with the essay all along, whereas those led to believe they had a low degree of choice reported no such agreement. What happened, apparently, was that people looking back after they had written the essay had no conscious memory of their pre-essay attitudes. It is as though the confabulation of intention erases its tracks, leaving people with no memory of ever having wanted other than what they currently see as their intention.

It is important to note that the success of cognitive dissonance and self-perception paradigms relies on the participants' impression that they are acting of their own choice. We do not confabulate reasons for our actions unless we also feel as though the actions were willed. For this reason we do not confabulate intentions for reactions that we understand don't require will, such as jerking a hand away from a hot stove. We know that such things will happen without our choice and do not originate from our desires or beliefs. It is only for actions that are willed that we apply a model of ideal agency and for which we need to have thoughts that precede our actions. Although the perception of choice has long been known as a key precondition for the phenomena of dissonance and self-perception (Wicklund & Brehm, 1976), it is only through an appreciation of these phenomena in the context of ideal agency that the role of this variable in making action seem willful can be understood. With action observed and will apparent, the person infers prior thoughts consistent with what was done.

Posthypnotic Suggestion

In the late 19th century, hypnotism reached the peak of its popularity. The public and the scientific community alike were fascinated with understanding hypnosis and its implications for medicine, psychoanalysis, and party games. Hypnosis also yields excellent illustrations of people confabulating intentions, as this is a common response to behaviors performed as the result of posthypnotic suggestion. Moll (1889) observed one such instance when he gave a woman the posthypnotic suggestion to take a book from the table and put it on the bookshelf when she woke up. He woke her, and she did as Moll had suggested. Moll asked her why she did this, and she reported that she did not like to see things untidy, and so she put the book away. She did not remember the suggestion that Moll had given her, and so she created a (very sensible) reason for her action in the form of a confabulated prior intention.

But hypnotism is no fun at all if you just make people put books on shelves. The woman in that case may well have been a tidy person and could have genuinely wanted to put that book away regardless of Moll's suggestion. The real power of hypnotism is that people can be persuaded to do very strange things that they might otherwise never do. It is easy to find a reason for a mundane action such as putting a book away, but it may not be so easy to find reasons for less sensible behaviors. Nevertheless, reasons can be found. In another case, Moll gave a man a posthypnotic suggestion to take a flowerpot from the window, wrap it in cloth, put it on the sofa, and bow to it three times. The man awoke and did as Moll had suggested. When asked why he did this, the man responded that the plant was probably cold. So, he wrapped it in cloth and put it near the fire (on the sofa) to keep it warm. This pleased him, so he bowed to the plant to show his happiness about the whole idea. He also added that this was not foolish, because he had reasons for doing it.

Left-Brain Interpreter

In the examples of posthypnotic suggestion and in the research on cognitive dissonance and self-perception, people confabulated a prior intention for their actions because they were unaware that the true cause for their action was created by others. Intention confabulation can also result when our true intentions originate within the self but remain concealed. Some fascinating examples of this emerge from "split brain" patients—people who have had the corpus callosum surgically severed as a treatment for severe epilepsy. The corpus callosum connects the left and right sides of the brain and is the means by which the two halves communicate with each other. As a result, people who have undergone this kind of procedure have two "minds," each partially unaware of what the other is thinking. Studies conducted by Gazzaniga (1983, 1988; Gazzaniga & LeDoux, 1978) have shown that if the intentions of the right half of the brain are concealed from the left, the person may confabulate an intention to match the action.

In these studies, the patient is presented with words or pictures to one visual field but not the other. Consequently, the information is only available to one half of the brain. The resulting actions must be quickly explained by the other side. For example, the patient J.W.'s left visual field (which corresponds to the right brain) was presented with the word *walk*. He obeyed and began to leave. When asked why, he responded that he was going to get a Coke. Prior testing had indicated that his right brain lacked the verbal sophistication to come up with such a response, so this answer must have been a quick interpretation of his behavior, provided by his left brain. In another instance, the word *laugh* was presented to his right brain, and J.W. laughed. When asked why he laughed, he replied, "You guys come up and test us every month. What a way to make a living" (Gazzaniga, 1983). Another explanation for laughter came when the patient N.G.'s right brain was presented with a nude picture of a woman. She laughed at the picture, but when asked why she did so, she responded that she thought the machine presenting the pictures was funny (Gazzaniga & LeDoux, 1978).

A telling example was found with the patient P.S., who was shown two pictures. One was shown to the left brain, and one to the right . Also in front of him was an array of pictures that were perceived by both visual fields. His task was to select the picture that best matched the picture he was shown. In one instance, a picture of a chicken was flashed to the left side of his brain at the same time that a snow scene was flashed to the right side of his brain. P.S. selected a picture of a claw with his right hand (left brain) to go with the picture of the chicken. He also selected a picture of a shovel with his left hand (right brain) to go with the snow scene. He was then asked to explain these choices. The left brain understood the choice of the claw; it matched the picture of the chicken it just saw. However, it did not know the reason for the choice of the shovel, which was selected by the right brain. It quickly improvised a reason for the choice—the shovel is used to clean out the chicken coop! The choice of the shovel was interpreted by the left brain in a way that was consistent with its own intentions and knowledge.

Gazzaniga (Gazzaniga & LeDoux, 1978) emphasized that the explanations that the patients give are offered as the reason for behavior, not merely as guesses at the reason. This is true despite the patients' awareness that their surgery has resulted in their having a "split brain." Even though they could say that it must be the result of their surgery, they are not inhibited from confabulating a different reason and feeling as though that reason is veridical. Gazzaniga (1988) accounted for these findings by suggesting that the left brain, in addition to its verbal duties, is responsible for interpreting our actions. The brain is made up of various modules that are responsible for specific functions and operate outside of our conscious awareness. Gazzaniga argued that there is a left-brain system (the interpreter) that monitors the actions that result from all these functional modules and comes up with the reasons for those actions. The interpreter considers all the outputs of the functional modules as soon as they are made and immediately constructs a hypothesis as to why a particular action occurred. In fact, the interpreter may not be privy to why a particular module responded. Nonetheless, it will take the behavior at face value and fit the event into the large outgoing mental schema (belief system) that it has already constructed.

According to Gazzaniga (Gazzaniga & LeDoux, 1978) this interpretation happens on a regular basis in the left side of the brain, which raises the question of how many of the intentions that we feel are ever the true reason for our actions. Even if people constantly confabulate their intentions, it is almost impossible to know because intentions are generally kept private. If we always had to make our intentions explicit before acting, like calling shots in a pool game, we might find a whole lot of actions that weren't as successful as we wanted and a whole lot of unexplained actions left over. As unpleasant as failure may be, having performed actions with no reason may be much more disconcerting. The left-brain interpreter provides a seamless connection between our thoughts and actions that makes us feel like effective, sensible, ideal agents.

INFERENCE OF THE WILL: APPARENT MENTAL CAUSATION

The feeling that we will many of our actions is an inescapable part of human life. We feel that we cause our actions when we think of going out to lunch and then do so, for example, and it is difficult in the face of this compelling feeling to subscribe to mechanistic theories of human behavior that suggest our actions are caused by psychological processes—not by our agent selves. The experience of willing an action is a variable one, however—something that is not inevitable in the action's occurrence—and this observation suggests that the feeling is separable from the processes whereby the action is caused. Quite simply, our "feeling of doing" any action is an inference, an interpretation that arises from specific conditions. The conditions that produced this feeling have been explored in the theory of apparent mental causation (Wegner, 2002; Wegner & Wheatley, 1999).

In essence, this theory says that we experience conscious will when we infer that our thought has caused our action. Thus, when you think of going to lunch and then do so, there's no telling what really caused the lunch going—but you nonetheless are likely to experience the strong sense that you consciously willed it. There are three principles guiding our inferences about the relationship between thought and action: priority, consistency, and exclusivity (Wegner 2002; Wegner & Wheatley, 1999). These principles suggest that we are most inclined to experience conscious will for an action when our thoughts immediately precede the action (priority), when these thoughts are consistent with the action (consistency), and when the thoughts occur exclusive of other salient potential cause of the action (exclusivity).

Thus, if you think that ice cream is tasty just before you lean over and have a bite of a cone, you are more likely to feel that you willed that action than if you think ice cream is tasty only after you take a bite (priority violated). You are more inclined to experience willing the action if you think that ice cream is tasty than if you think it is unhealthy when you take that bite (consistency violated). And you are more likely to think you willed the bite if your thought accompanies the action alone—and is not accompanied by the ministrations of a pushy friend who is encouraging you to try the cone (exclusivity violated). To the extent that thoughts about an action have all these aspects, the thoughts provide a preview of the action of the agent, and it feels as though the will has operated by transforming the goals of the agent into reality.

Evidence on the Principles

Evidence for apparent mental causation comes from studies that demonstrate that giving people a preview of action does increase their feelings of authorship for that action. In a study inspired by the Ouija board, the role of the priority of thought was investigated (Wegner & Wheatley, 1999). A participant and a confederate were seated facing each other across a small table, both wearing headphones. Between them was a computer mouse and on a computer screen beside

them was a page from the book *I Spy*, displaying about 50 small objects, such as a swan, a dinosaur, a car, and the like. The participant and the confederate were told that they were to move the mouse together by both placing their fingertips on it, and that they should stop moving the mouse every 30 s or so when music played. When they made a stop, they rated how intentional each stop was—how much they personally intended to make the stop independent of their partner's actions. They were also played words through the headphones during the study, which was ostensibly to provide some distraction, and were told that their partner would hear different words than they would.

All this was designed to create a situation in which the participant would have a thought about the action before it was actually performed. Moreover, the action would not be his or her own. The confederate was instructed to guide the motion of the mouse on some trials and to make specific stops on certain objects (e.g., the swan). Some of the words that the participants heard corresponded to the objects on the screen (e.g., *swan*). Each of the stops was timed to occur at specific intervals from when the participant heard the corresponding word, so the participant was primed with a thought of the action either 30 s before the stop, 5 s before, 1 s before, or 1 s after. Participants rated the stop as much more intentional when they heard the prime either 5 s before or 1 s before the stop. When the prime was given 30 s before or 1 s after, they felt the stop was less intentional on their part. When the thought and the action were consistent in a timely fashion, people came to feel that the action was their own choice, even though in all cases the stops were controlled by the confederate.

Part of the reason why magic tricks and optical illusions capture the fascination of eager audiences is that they persist despite the watcher's knowing that they can't be so. If apparent mental causation really is an illusion, then the feeling of conscious will could be induced even in cases in which people logically know that they did not create the action. Wegner, Sparrow, and Winerman (2004) looked at just this. People in a lab setting rated how much control and intention they felt about moving someone else's arms, but under conditions that paralleled real self-generated action. To promote the illusion, the experimenters made it look as though the other person's arms were the participant's own arms. Participants watched themselves in a mirror as another person hidden immediately behind them extended arms forward on each side of them. This is the pantomime routine sometimes known as *helping hands*. The person behind the participant then followed a series of instructions for hand movement delivered over headphones (e.g., "clap three times"). In some cases, the participant heard these same instructions and so knew what actions were to follow. In other cases, the participant did not hear any instructions. Afterward, participants were asked to rate on a 7-point scale how much control they felt over the other person's arms. The people who heard the previews felt a greater degree of control over the arms than those who did not, even though they knew that they were not actually controlling any of the movement.

Research has also examined how apparent mental causation can give rise to magical thinking (Pronin & Wegner, 2003). In two studies, participants were led to have evil thoughts about another person, who then came to experience a

negative outcome. Would participants feel as though their ill thoughts produced an ill will? Participants were told they were participating in a study on voodoo with another person (actually a confederate). The participant was selected to play the role of witch doctor, and the confederate played the role of the victim. The participant and confederate went to separate rooms, and the participant said a brief voodoo chant (written by the experimenter, not a genuine chant). In one condition, the chant was neutral; in the other condition, the chant was malevolent and specifically wished harm against the confederate. After the chant, the participant was instructed to stick pins in a voodoo doll (with the confederate's name on it) in the presence of the confederate. Following the voodoo ritual, the confederate reported that she now had a headache that she did not have before. Participants who were instructed to recite the malevolent chant felt more responsible for the confederate's headache than those who just said a neutral chant. In a second study, negative thoughts about the confederate were manipulated by having the confederate act in an obnoxious and irritating manner. In this condition the confederate arrived 10 min late to the study, littered on the floor, chewed gum with his mouth open, and wore a T-shirt that read "Stupid people shouldn't breed." In a neutral condition the confederate acted normally and wore an inoffensive shirt. Again, the participant was selected to be the witch doctor and the confederate was the victim. Rather than say a chant, the participant was instructed to focus his or her thoughts on the confederate. As expected, participants reported more negative thoughts toward the confederate when he acted offensively and also felt more responsible for the confederate's feigned headache.

Illusion of Control

Evidence for apparent mental causation also comes from Langer's (1975; Langer & Roth, 1975) work on the illusion of control. She observed that people often feel that they have personal control over the outcome of events in which control is impossible. For example, they may feel that a game of chance (e.g., throwing dice) is actually a game of skill. However, the illusion of control does not arise universally and equally for all events. Certain factors such as choice, familiarity, and involvement may increase feelings of control in cases where no control exists. Performing the actions one normally performs in cases of skill may, in chance tasks, lead to the illusory sense that one actually can control the outcome.

In one study (Langer, 1975) participants bought raffle tickets for $1. Half the participants were allowed to choose which ticket they wanted; others were simply given a ticket. Familiarity was also manipulated. Half the tickets had a letter of the alphabet on it; the other half had a novel symbol. Participants were later called and asked whether they would be willing to swap the ticket they bought for a ticket in another lottery, one with the same prize but with better odds. People were more attached to their ticket if they had selected the ticket themselves—despite the better odds, they were less likely to trade their ticket. Familiarity also affected the decision to swap; people were less inclined to trade if it had a letter of the alphabet than a novel symbol. Another study examined the role of involvement. People at a racetrack were entered into a lottery as part of their

admission fee. They were approached throughout the day and asked about the likelihood that they would win. Confidence increased as time went by. Langer reasoned that this was a result of the time that people had to think about winning. With more time, they had more opportunity to think about their "strategies" for winning.

The illusion of control is fostered when one behaves in a way that resembles what one would do in a skill situation. Thus, just acting like one has control has the side effect of feeling that one is in control. Becoming involved in the task by choosing a card or some other action gives a person opportunity to think about winning, even if by a random action. Such thoughts promote apparent mental causation. Familiarity and time to think about the event can increase intentionality because they direct a person's thoughts toward the event. Thoughts about action make people feel as though they are consciously willing the desired outcome, even when influence over that event is impossible.

For the ideal agent, will is the servant of thought. If the will of the agent does not reflect the thoughts and intentions of the agent, something has gone wrong. Thus, when one feels the intention to have a certain outcome, one also feels that one is doing things that can control that outcome. This association is strong even in cases in which actual control is not possible. The more one thinks about an outcome, the more one feels that outcome is within personal control when it happens. Thought and action lead to an inference that the action was willed.

INFERENCE OF ACTION: ACTION MISPERCEPTION

During his campaign for the U.S. presidency, then Vice President Al Gore was interviewed on the CNN program *Late Edition.* In that interview he made his famous claim that he "took the initiative in creating the Internet." The statement was altered in the retelling, and soon his claim to have "invented the Internet" was being ridiculed by the press and public. Some people who opposed the vice president's campaign called him a liar, and those who were more sympathetic suggested that he was merely exaggerating. It is true that while in the Senate, Gore had introduced bills that promoted the growth of the Internet, but did his actions actually result in the Internet as we know it today? Perhaps Gore was neither lying nor boasting but simply expressing his pursuit of ideal agency. He perceived his actions to be somewhat grander and more successful than they were.

This may be easy to understand if we assume that Gore had both thought about seeing the Internet develop and was willfully attempting to contribute to that development. For the ideal agent, action is the culmination of thought and will. Thought and will should naturally lead to a consistent action and, most important, to an outcome that matches the original intention reflected in the person's thoughts. To an outside observer who does not remember Al Gore wanting or trying to develop the Internet, his claim seems like an outrageous leap. However, to Gore, his intention and will were consistent with creating the Internet—and it happened! It is a relatively small step to feel that his actions did produce the outcome.

Such misperception can also happen when there is not even an outcome to claim. A person playing tennis could "see" the ball fall inside the line when everyone else sees it fall outside, for instance, or a person telling a joke could think it was a hit when everyone else thought it was a flop. When our best plans go awry, we may feel that the result was more successful than it was, so much so that we misperceive the actual outcome. This is a matter of action misperception, and it happens when we revise our actions to be consistent with our thoughts and our will.

Because actions are less open to interpretation than are thought or will, misperceiving action may be a more difficult maneuver for an agent than either intention confabulation or apparent mental causation. Yet, there is evidence that people distort what they think of their behavior when information about the action is somewhat ambiguous. For example, people who are incompetent in humor, grammar, or logic tend to overestimate their own skills in those respective domains (Kruger & Dunning, 1999). Children with poor social skills seem unaware of their difficulty and judge themselves as socially competent (Fagot & O'Brien, 1994) as do adults in the same predicament (Bem & Lord, 1979). Research on the "above-average effect" shows that people judge themselves to be better than most other people on many different attributes, a bias that increases when the attribute itself is loosely defined (Dunning, Meyerowitz, & Holzberg, 1989).

The ambiguity of an action increases as the action is further away in time, so past actions are more susceptible to misperception. In a classic demonstration of memory distortion, Hastorf and Cantril (1954) found that people who watched the same college football game remembered the game differently depending on which college they went to, judging the opposing team to be playing less fairly than their own. In a related study, Ross and Sicoly (1979) found that husbands and wives claimed responsibility for a greater proportion of the housework than they attributed to their spouse, accounting for more than 100% between the couples. Such distortions tend to be of the self-serving variety and so can be explained by the need to maintain self-esteem (Tesser, 1988). In these various cases, people may recall that they wanted to be funny, work hard, be smart or socially competent, be good at tests, and have other positive attributes, so they perceive themselves to be successful at these efforts. Although it remains to be seen whether such effects indeed become more pronounced when both prior thought and feelings of will are present, it makes sense that these action misperceptions occur as a result of strivings for ideal agency.

Dissonance Revisited

Most cognitive dissonance studies focus on attitude change. However, it may not always be possible to change a prior attitude if the person is acutely aware of this prior attitude. People are less free to alter their intentions and prior thoughts when these are explicit beforehand—at least not without recognizing that it is indeed an alteration. If the memory of prior thoughts and a prior attitude cannot be changed, other cognitions could be changed, including cognitions about their actual behavior. In one study, Scheier and Carver (1980) made people's prior

attitude salient to them through self-attention. Half the participants were put in front of a mirror, which directed the participants' thoughts toward themselves. Other participants were not put in this self-focused situation. The researchers then had participants write a counterattitudinal essay. When prior attitudes were made salient in this way, participants did not revise their prior attitudes to match their essay. Rather, they distorted what they thought of the essay they wrote. Participants who were initially made to focus on themselves later judged the counterattitudinal essay they penned as being closer to their own attitude. However, this judgment was inaccurate; raters blind to condition did not think the essays reflected the original attitude. Action misperception occurred to bring the perceived action into line with prior thoughts.

Action Misperception Studies

Although there is considerable evidence that people misperceive their own actions in a self-flattering way when they are given the chance, the role of intention and conscious will in that misperception has not been specifically examined. Preston and Wegner (2003) conducted two studies to investigate this relationship. In a first study, we looked at the effect of exerting will over an action. Participants came into the lab to do a study on typing. A seven-letter word appeared on the computer screen before them, and their task was to type the word as accurately as possible without looking at the keys. To prevent them from cheating, the letters on the keyboard were covered. The degree of conscious will people felt over this act was manipulated by giving some degree of choice over what to type. On some trials, there was no choice—one word was presented that the participant had to type. On other trials, two words were presented side by side and participants selected which one they wanted to type. After typing each word, participants had to report how many of the letters they thought they had typed correctly. Participants felt they had done a better typing job when they were given a chance to exert their will; they thought they had typed more letters correctly when they were given a choice of word to type, as compared to the trials in which they had no choice.

In another study, Preston and Wegner (2003) manipulated both thoughts about action and the experience of will. Participants fired a toy gun at a target 10 ft (approximately 3 m) away. The gun shot a foam bullet that hit the target and fell to the ground. The task of the participants was to judge how close (in inches) they were to hitting the bull's-eye of the target after each shot. The participants could not see it from where they were standing, but the target had a faint grid of 1-in. (approximately 2.6-cm) squares, allowing the experimenter surreptitiously to record the actual distance of each shot from the bull's-eye. Will was manipulated by the type of firing instruction that the participant received before each shot. On some trials, the experimenter counted down from three before instructing the participant to fire ("3–2–1–Go!"); on other trials, the instruction to fire came with no countdown (Go!). Pretests revealed that giving people a countdown before an action increases the degree to which the action feels personally willed.

The intention to be accurate was manipulated by what the target represented. Slides of famous faces were projected onto the target, some of which were pictures of people who were well-liked (e.g., Mahatma Gandhi) and others of which were of people who were disliked (e.g., Adolf Hitler) with the assumption that people would be more motivated to hit a person they despised. The bull's-eye was centered directly between the eyes of each famous face. Both motivation to be accurate (i.e., shooting at someone they hate) and the conscious will to fire contributed to people's flattering estimates of their own accuracy. Controlling for actual distance from the bull's-eye, people judged themselves to be most accurate in cases in which they were shooting at someone they disliked.

These studies provide some hints about the possible influence that thought and will may have on action misperception. It may well be that many of the most brazen, self-aggrandizing judgments of action occur not because people are generally egomaniacal. Instead, the tendency to misperceive action may follow primarily when people think about an action in advance of its occurrence and feel they have willed that action. Then, it begins to make sense that the action, too, should fall into line. We only start distorting our view of what we have done when we have thought about doing it and feel that we chose to do it.

CONCLUSION

All agents are not created equal. The hallmark of a good agent is its ability to manipulate the environment in a self-serving way. Such manipulation in humans involves three components—thought, will, and action. Having thoughts before acting allows an agent to sense the features of the environment relevant to action and so to select more favorable and more complex actions. The experience of will allows the agent to feel "in charge" of the specific actions and attribute agency for the action to the self. And of course, if the actions themselves are successful, the agent is successful as well. An ideal agent possesses all of these characteristics working together—thoughts create a goal for the agent to achieve, the agent experiences willing itself to pursue this goal, and its actions then satisfy the goal. This is what we aspire to be in our own agency, emulating this model for many of our actions each day.

These components of agency are not always easy to see. Action is observable some of the time, of course, but thoughts occur and then disappear, and the will is simply an experience that occurs on the way from thought to action. The model of an ideal agent leads us to believe that when we start to make out the image of an agent in some way—through thoughts, will, or action—the complete picture may soon come into view. We go beyond the information given and infer the remaining facets of ideal agency.

Anyone who has spent much time as a human knows quite well that ideal agency is not always possible. When we fall short of the ideal, we infer that the missing components of ideal agency are actually in place. As we have seen, such distortions take three forms: intention confabulation (when we have will and action but must infer thought), apparent mental causation (when we have thought and

action and so infer will), and action misperception (when we have will and thought but must infer action). The evidence we have presented suggests that people can distort their perceptions in these three ways in order to maintain a conception of self as an ideal agent.

Ultimately, of course, all of these processes lead us away from reality; they involve changes in self-perception that may carry the person far from a veridical view. At the extreme, these sources of error and illusion in self-understanding may be responsible for psychological disorders and maladaptive behavior. A person who always insisted that every action was premeditated, who believed that every thought about action that happened to predict the action meant that the action was willed, or who perceived all willed actions as successful, after all, would qualify for a variety of psychiatric diagnoses. For the normal person, ideal agency guides inferences about the self within the constraints provided by reality.

ACKNOWLEDGMENTS

This chapter was prepared with the support of National Institute of Mental Health Grant MH 49127. We thank Ricardo Alanis, Marie Baldwin, and Michael Berkovits for their assistance in data collection.

REFERENCES

Abbott, K., & Flavell, J. H. (1996). *Young children's understanding of intention.* Unpublished manuscript, Stanford University, Stanford, CA.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Anscombe, G. E. M. (1957). *Intention.* London: Blackwell.

Arkin, R. M., & Baumgardner, A. H. (1985). Self-handicapping. In J. H. Harvey & G. W. Weary (Eds.), *Attribution: Basic issues and applications* (pp. 169–202). Orlando, FL: Academic Press.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive, 13,* 513–552.

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences, 4,* 29–34.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review, 74,* 183–200.

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York: Academic Press.

Bem, D. J., & Lord, C. G. (1979). Template matching: A proposal for probing the ecological validity of experimental settings in social psychology. *Journal of Personality and Social Psychology, 37,* 833–846.

Bem, D. J., & McConnell, H. K. (1970). Testing the self-perception explanation of dissonance phenomena: On the salience of premanipulation attitudes. *Journal of Personality and Social Psychology 14,* 23–31.

Berglas, S., & Jones, E. E. (1978). Drug choice as a self-handicapping strategy in response to noncontingent success. *Journal of Personality and Social Psychology, 36,* 405–417.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews: Neuroscience, 2,* 561–567.

Boyer, P. (2001). *Religion explained.* New York: Basic Books.

Brand, M. (1984). *Intending and acting.* Cambridge, MA: MIT Press.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior.* Cambridge, England: Cambridge University Press.

Cohen, A. R. (1962). An experiment on small rewards for discrepant compliance and attitude change. In J. W. Brehm & A. R. Cohen (Eds.), *Explorations in cognitive dissonance* (pp. 73–78). New York: Wiley.

Cooley, C. H. (1902). *Human nature and the social order.* New York: Scribners.

Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 229–266). San Diego, CA: Academic Press.

Davidson, D. (1980). Agency. In *Actions and events* (pp. 43–62). Oxford, England: Clarendon Press.

DeCharms, R. (1968). *Personal causation.* New York: Academic Press.

Descartes, R. (1901/1989). *Discourse on method and mediations* (J. Veitch, Trans.). New York: Prometheus. (Original work published 1641)

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57,* 1082–1090.

Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review, 95,* 256–273.

Fagot, B. I., & O'Brien, M. (1994). Activity level in young children: Cross-age stability, situational influence, correlates with temperament, and perception of problem behaviors. *Merrill-Palmer Quarterly, 40,* 378–398.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58,* 203–210.

Frith, C., Blakemore, S. J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London, Series B, 355,* 1771–1788.

Gazzaniga, M. S. (1983). Right hemisphere language following brain bisection: A 20-year perspective. *American Psychologist 38,* 525–537.

Gazzaniga, M. S. (1988). Brain modularity: Towards a philosophy of conscious experience. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 218–238). Oxford, England: Clarendon Press.

Gazzaniga, M. S., & LeDoux, J. E. (1978). *The integrated mind.* New York: Plenum Press.

Guthrie, S. E. (1980). A cognitive theory of religion. *Current Anthropology, 21,* 181–203.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49,* 129–134.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57,* 243–259.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*, 1121–1134.

Langer, E., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a pure chance task. *Journal of Personality and Social Psychology, 32*, 951–955.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311–328.

Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology, 6*, 245–254.

Moll, A. (1889). *Hypnotism*. London: Walter Scott.

Newell, A. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 463–526). New York: Academic Press.

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

Powers, W. T. (1990). Control theory: A model of organisms. *System Dynamics Review, 6*, 1–20.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral & Brain Sciences, 1*, 515–526.

Preston, J., & Wegner, D. M. (2003). *Action misperception as a consequence of intention and will to act*. Unpublished manuscript, Harvard University, Cambridge, MA.

Pronin, E., Wegner, D. M., & McCarthy, K. (2004). *Everyday magical powers: The role of apparent mental causation in the overestimation of personal influences*. Unpublished manuscript, Harvard University, Cambridge, MA.

Rogers, C. R. (1961). *On becoming a person*. Boston: Houghton Mifflin.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology, 37*, 322–336.

Russell, S. J., & Norvig, P. (2003). *Artificial intelligence: A modern approach* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Scheier, M. F., & Carver, C. S. (1980). Private and public self-attention, resistance to change, and dissonance reduction. *Journal of Personality and Social Psychology, 39*, 390-405.

Schult, C. A. (1997). *Intended actions and intentional states: Young children's understanding of the causes of human actions*. (Doctoral dissertation, University of Michigan, 1997). *Dissertation Abstracts International, Section B: The Sciences and Engineering, 57*, 11-B.

Shapiro, R., & Gilbert Abramson, R. (Producers), & Burton, T. (Director). (1985). *Pee-Wee's big adventure* [Motion picture]. United States: Warner Brothers.

Spiegel, L. (Executive Producer). (1999, March 9). *Late edition* [Television broadcast]. Washington, DC: Cable News Network.

Stephens, G. L., & Graham, G. (2000). *When self-consciousness breaks: Alien voices and inserted thoughts*. Cambridge, MA: MIT Press.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowiz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181–227). San Diego, CA: Academic Press.

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

Wegner, D. M., & Sparrow, B. (in press). Authorship processing. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (3rd ed.). 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, 86*, 838–848.

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

Wegner, D. M., Winerman, L., & Sparrow, B. (2003). *Action at a distance: Experiencing control over the movements of others*. Manuscript submitted for publication.

Wellman, H. M. (1992). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Hillsdale, NJ: Erlbaum.

Wilson, T. D., Hodges, S. D., & LaFleur, S. J. (1995). Effects of introspecting about reasons: Inferring attitudes from accessible thoughts. *Journal of Personality and Social Psychology, 69*, 16–28.

Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*, 181–192.