BEHAVIORAL AND BRAIN SCIENCES (2004) **27,** 649–692
Printed in the United States of America

Précis of *The illusion of conscious will*

Daniel M. Wegner
Department of Psychology, Harvard University, Cambridge, MA 02138.
wegner@wjh.harvard.edu **http://www.wjh.harvard.edu/~wegner/**

Abstract: The experience of conscious will is the feeling that we are doing things. This feeling occurs for many things we do, conveying to us again and again the sense that we consciously cause our actions. But the feeling may not be a true reading of what is happening in our minds, brains, and bodies as our actions are produced. The feeling of conscious will can be fooled. This happens in clinical disorders such as alien hand syndrome, dissociative identity disorder, and schizophrenic auditory hallucinations. And in people without disorders, phenomena such as hypnosis, automatic writing, Ouija board spelling, water dowsing, facilitated communication, speaking in tongues, spirit possession, and trance channeling also illustrate anomalies of will – cases when actions occur without will or will occurs without action. This book brings these cases together with research evidence from laboratories in psychology to explore a theory of *apparent mental causation.* According to this theory, when a thought appears in consciousness just prior to an action, is consistent with the action, and appears exclusive of salient alternative causes of the action, we experience conscious will and ascribe authorship to ourselves for the action. Experiences of conscious will thus arise from processes whereby the mind interprets itself – not from processes whereby mind creates action. Conscious will, in this view, is an indication that we *think* we have caused an action, not a revelation of the causal sequence by which the action was produced.

Keywords: apparent mental causation; automatism; conscious will; determinism; free will; perceived control

1. The illusion (Ch. 1)

So, here you are, reading about conscious will. How could this have happened? One way to explain it would be to examine the causes of your behavior. A team of scientists could study your reported thoughts, emotions, and motives, your genetics and your history of learning, experience, and development, your social situation and culture, your memories and reaction times, your physiology and neuroanatomy, and lots of other things as well. If they somehow had access to all the information they could ever want, the assumption of psychology is that they could uncover the mechanisms that give rise to all your behavior, and so could certainly explain why you are reading these words at this moment. However, another way to explain the fact of your reading these lines is just to say that you decided to begin reading. You consciously willed what you are doing.

The ideas of conscious will and psychological mechanism have an oil and water relationship, having never been properly reconciled. One way to put them together is to say that the mechanistic approach is the explanation preferred for scientific purposes, but that the person's experience of conscious will is utterly convincing and important to the person – and so must be understood scientifically as well. The mechanisms underlying the experience of will are themselves a fundamental topic of scientific study.

1.1. Conscious will

Conscious will is usually understood in one of two ways. It is common to talk about conscious will as something that is experienced when we perform an action: Actions feel willed or not, and this feeling of voluntariness or doing a thing "on purpose" is an indication of conscious will. It is also common, however, to speak of conscious will as a force of mind, a name for the causal link between our minds and our actions. One might assume that the *experience* of consciously willing an action and the *causation* of the action by the person's conscious mind are the same thing. As it turns out, however, they are entirely distinct, and the tendency to confuse them is the source of the illusion of conscious will. So, to begin, we will need to look into each in turn, first examining will as an experience and then considering will as a causal force.

1.1.1. The experience of conscious will. Will is a feeling. David Hume was sufficiently impressed by this idea that he proposed to define the will as "nothing but *the internal impression we feel and are conscious of, when we knowingly give rise to any new motion of our body, or new perception of our mind*" (Hume 1739/1888, p. 399, emphasis in origi-

DANIEL M. WEGNER is Professor of Psychology at Harvard University. Since his 1974 Ph.D. from Michigan State University, he has published six books (including *White Bears and Other Unwanted Thoughts*) and more than a hundred articles. His research on thought suppression, mental control, action identification, transactive memory, and conscious will has been funded by the National Science Foundation and by the National Institute of Mental Health. A 1996–1997 Fellow of the Center for Advanced Study in the Behavioral Sciences, he has served as Associate Editor of *Psychological Review* and currently is on the Board of Reviewing Editors of *Science.*

 0140-525X/04 $12.50

https://doi.org/10.1017/S0140525X04000159 Published online by Cambridge University Press

nal). This definition puts the person's experience at the very center of the whole concept – the will is not some cause or force or motor in a person, but rather is the personal conscious feeling of such causing, forcing, or motoring. Hume's definition makes sense because the occurrence of this conscious experience is an absolute *must* for anyone to claim to have done something that he or she consciously willed.

Without an experience of willing, even actions that look entirely voluntary from the outside still fall short of qualifying as truly *willed.* Intentions, plans, and other thoughts can be experienced, and still the action is not willed if the person says it was not. If a person plans to take a shower, for example, and says that she intends to do it as she climbs into the water, spends 15 minutes in there scrubbing up nicely, and then comes out reporting that she indeed seems to have had a shower – but yet also reports *not* feeling she had consciously willed her showering – who are we to protest? Consciously willing an action requires a feeling of doing (Ansfield & Wegner 1996), a kind of internal "oomph" that somehow certifies authentically that one has done the action. If the person did not get that feeling about her shower, then even if we climbed in with her to investigate, there is no way we could establish for sure whether she consciously willed her showering.

The fact that experiences of conscious will can only be established by self-reports ("I showered, yes I did") would be quite all right if the self-reports always corresponded with some other outward indication of the experience. However, this correspondence does not always happen. The experience of will that is so essential for the occurrence of consciously willed action does not always accompany actions that appear by other indications to be willed.

Consider, for example, the case of people who have *alien hand syndrome,* a neuropsychological disorder in which a person experiences one hand as operating with a mind of its own. Alien hand patients typically experience one hand as acting autonomously. They do not experience willing its actions, and may find it moving at cross-purposes with their conscious intention. This syndrome is often linked with damage to the middle of the frontal lobe on the side of the brain opposite the affected hand (Gasquoine 1993). Banks and colleagues (1989) report an alien hand patient whose

> left hand would tenaciously grope for and grasp any nearby object, pick and pull at her clothes, and even grasp her throat during sleep. . . . She slept with the arm tied to prevent nocturnal misbehavior. She never denied that her left arm and hand belonged to her, although she did refer to her limb as though it were an autonomous entity. (Banks et al. 1989, p. 456)

Should the alien hand's movements be classed as willed or unwilled? On the one hand (pun can't be helped), the alien hand seems to do some fairly complicated things, acts we might class as willful and voluntary if we were just watching and hadn't learned of the patient's lamentable loss of control. In the case of another patient, for example,

> While playing checkers on one occasion, the left hand made a move he did not wish to make, and he corrected the move with the right hand; however, the left hand, to the patient's frustration, repeated the false move. On other occasions, he turned the pages of the book with one hand while the other tried to close it; he shaved with the right hand while the left one unzipped his jacket; he tried to soap a washcloth while the left hand kept putting the soap back in the dish; and he tried to open a closet with the right hand while the left one closed it. (Banks et al. 1989, p. 457)

By the looks of it, the alien hand is quite willful. On the other hand (as the pun drags on), however, the patient does not experience these actions as consciously willed.

Brain damage is not the only way that the experience of will can be undermined. Consider, for instance, the feelings of involuntariness that occur during hypnosis. Perhaps the most profound single effect of hypnosis is the feeling that your actions are happening to you, rather than that you are doing them (Lynn et al. 1990). To produce this experience, a hypnotist might suggest, "Please hold your arm out to your side. Now, concentrate on the feelings in your arm. What you will find is that your arm is becoming heavy. It feels as though a great weight were pulling it down. It is so very heavy. It is being pulled down, down toward the ground. Your arm is heavy, very heavy. It is getting so heavy you can't resist. Your arm is falling, falling down toward the ground." With enough of this patter, many listeners will indeed experience the arm becoming heavy, and some will even find their arm falling down. When quizzed on it, these individuals often report that they felt no sense of moving their arm voluntarily, but rather experienced the downward movement as something that happened to them. This does not occur for everyone in this situation, only for some, but it nonetheless indicates that the experience of will can be manipulated in a voluntary action.

In the case of hypnotic involuntariness, the person has a very clear and well-rehearsed idea of the upcoming action. Admittedly, this idea of the action is really phrased more as an expectation ("My arm will fall") than as an intention ("I will lower my arm"), but it nonetheless occurs before the action when an intention normally happens, and it provides a distinct preview of the action that is to come (Kirsch & Lynn 1998b; Spanos 1986b). Hypnotic involuntariness thus provides an example of the lack of experience of will that is yet more perplexing than alien hand syndrome. With alien hand, the person simply does not know what the hand will do, but with hypnosis, conscious will is lacking – even when knowledge of the action is present. And without the *experience* of willing, even this foreknowledge of the action seems insufficient to move the action into the "consciously willed" category. If it does not feel as though you did it, then it does not seem that the will was operating.

Another case of the absence of an experience of will occurs in *table-turning,* a curious phenomenon discovered in the spiritualist movement in Europe and the United States in the mid-nineteenth century (Ansfield & Wegner 1996; Carpenter 1888). To create this effect, a group of people sits around a table with their hands on its surface. If they are convinced that the table might move as the result of spirit intervention (or if they are even just hoping for such an effect), and sit patiently waiting for such movement, it is often found that the table *does* start to move after some time. It might even move about the room or begin rotating so quickly that the participants can barely keep up. Carpenter (1888, pp. 292–93) observed "all this is done, not merely without the least consciousness on the part of the performers that they are exercising any force of their own, but for the most part under the full conviction that they are not." Incidentally, table-turning was sufficiently controversial that it attracted the attention of the chemist and physicist Michael Faraday, who proceeded to test the source of the table movement. He placed force measurement devices between participants' hands and the table, and found that the

Table 1. *Conditions of human action*

	Feeling of Doing	No Feeling of Doing
Doing	Normal voluntary action	Automatism
Not Doing	Illusion of Control	Normal inaction

source of the movement was their hands and not the table (Faraday 1853).

Such examples of the separation of action from the experience of will suggest that it is useful to draw a distinction between them. Table 1 shows four basic conditions of human action – the combinations that arise when we emphasize the distinction between action and the sense of acting willfully. The upper left corner contains the expected correspondence of action and the feeling of doing – the case when we do something and feel also that we are doing it. This is the noncontroversial case, or perhaps the assumed human condition. The lower right corner is also noncontroversial, the instance when we are not doing anything and feel we are not.

The upper right – the case of no feeling of will when there is in fact the occurrence of action – encompasses the examples we have been inspecting thus far. The movement of alien hands, the case of hypnotic suggestion of arm heaviness, and table-turning all fit this quadrant, as they involve no feeling of doing in what appear otherwise to be voluntary actions. These can be classed in general as *automatisms.* The other special quadrant of the table includes cases of the *illusion of control.* Ellen Langer (1975) used this term to describe instances when people have the feeling that they are doing something when they actually are not doing anything.

The illusion of control is acute in our interactions with machines – as when we do not know whether our push of an elevator button or Coke machine selection has done anything, yet sense that it has. The illusion is usually studied with judgments of contingency (e.g., Matute 1996) by having people try to tell whether they are causing a particular effect, for example, turning on a light, by doing something, such as pushing a button, when the button and the light are not perfectly connected and the light may flash randomly by itself. But we experience the illusion, too, when we roll dice or flip coins in a certain way, hoping that we will thus be able to influence the outcome. It even happens sometimes that we feel we have contributed to the outcome of a sporting event on TV just by our presence in the room ("Did I just jinx them by running off to the fridge?").

Most of the things we do in everyday life seem to fall along the "normal" diagonal in this fourfold table. Action and the experience of will usually correspond, so we feel we are doing things willfully when we actually do them, and feel we are not doing something when in truth we have not done it. Still, the automatisms and illusions of control that lie off this diagonal remind us that action and the feeling of doing are not locked together inevitably. They come apart often enough that one wonders whether they may be produced by separate systems in the mind. The processes of mind that produce the experience of will may be quite distinct from the processes of mind that produce the action itself. As soon as we accept the idea that the will should be understood as an *experience* of the person who acts, we come to realize that conscious will is not inherent in action – there are actions that have it and actions that do not.

1.1.2. The force of conscious will. Will is not only an experience, but also a force. Because of this, it is tempting to think that the conscious experience of will is a direct perception of the force of will. The *feeling* that one is purposefully not having a cookie, for example, can easily be taken as an immediate perception of one's conscious mind *causing* this act of self-control. We seem to experience the force within us that keeps the cookie out of our mouths, but the force is not the same thing as the experience.

When conscious will is described as a force, it can take different forms. Will can come in little dabs to produce individual acts, or it can be a more long-lasting property of a person, a kind of inner strength or resolve. Just as a dish might have hotness or an automobile might have the property of being red, a person seems to have will, a quality of power that causes his or her actions. The force may be with us. Such will can be strong or weak, and so can serve to explain things such as one person's steely persistence in the attempt to dig a swimming pool in the back yard, for example, or another person's knee-buckling weakness for chocolate. The notion of strength of will has been an important intuitive explanation of human behavior since the ancients (Charlton 1988), and it has served throughout the history of psychology as the centerpiece of the psychology of will. The classic partition of the mind into three functions includes cognition, emotion, and *conation* – the will or volitional component (e.g., James 1890).

The will in this traditional way of thinking is an explanatory entity of the first order. In other words, it explains lots of things but nothing explains it. As Joseph Buchanan (1812) described it, "Volition has commonly been considered by metaphysical writers, as consisting in the exertion of an innate power, or constituent faculty of the mind, denominated will, concerning whose intrinsic nature it is fruitless and unnecessary to inquire" (p. 298). At the extreme, of course, this view of the will makes the scientific study of it entirely out of the question, and suggests instead that it ought to be worshiped. Pointing to will as a force in a person that causes the person's action is the same kind of explanation as saying that God has caused an event. This is a stopper that trumps any other explanation, but that still seems not to explain anything at all in a predictive sense. Just as we cannot tell what God is going to do, we cannot predict what the will is likely to do either.

The notion that will is a force residing in a person has a further problem. Hume remarked on this when he described the basic difficulty that occurs whenever a person perceives causality in an object. Essentially, he pointed out that causality is not a property inhering in objects. For instance, when we see a bowling ball go scooting down the lane and smashing into the pins, it certainly *seems* as though the ball has some kind of causal force in it. The ball is the cause and the explosive reaction of the pins is the effect. Hume pointed out, though, that you cannot *see* causation in something, but must only infer it from the constant relation between cause and effect. Every time the ball rolls into the pins, they bounce away. Ergo, the ball caused the pins to move. But there is no property of causality nestled somewhere in that ball, or hanging somewhere in space between the ball and pins, that somehow works this magic. Causation is an event, not a thing or a characteristic or attribute of an object.

In the same sense, causation cannot be a property of a person's conscious intention. You can't *see* your conscious intention causing an action, but can only infer this from the regular relation between intention and action. Normally, when you intend things, they happen. Hume remarked in *A Treatise on Human Nature* (1739/1888) that the "constant union" and "inference of the mind" that establishes causality in physical events must also give rise to causality in "actions of the mind." He said:

> Some have asserted . . . that we feel an energy, or power, in our own mind. . . . But to convince us how fallacious this reasoning is, we need only consider . . . that the will being here consider'd as a cause, has no more a discoverable connexion with its effects, than any material cause has with its proper effect. . . . In short, the actions of the mind are, in this respect, the same with those of matter. We perceive only their constant conjunction; nor can we ever reason beyond it. No internal impression has an apparent energy, more than external objects have. (pp. 400–401)

Hume realized, then, that calling the will a force in a person's consciousness – even in one's own consciousness – must always overreach what we can see (or even introspect), and so should be understood as an attribution or inference.

This is not to say that the concept of will power is useless. Rather, Hume's analysis suggests that the concepts of force of will or will power must be accompanied by careful causal inference. These ideas can be used as the basis for scientific theories of human behavior, certainly, as they serve as summaries of the degree of relationship that may exist between the mind and behavior. But we must be careful to distinguish between such *empirical will* – the causality of the person's conscious thoughts as established by a scientific analysis of their covariation with the person's behavior – and the *phenomenal will* – the person's reported experience of will. The empirical will can be measured by examining the degree of covariation between the person's self-reported conscious thought and the person's action, and by assessing the causal role of that thought in the context of other possible causes of the action (and possible causes of the thought as well).

The empirical will – the actual relationship between mind and action – is a central topic of scientific psychology. In psychology, clear indications of the empirical will can be found whenever causal relationships are observed between people's thoughts, beliefs, intentions, plans, or other conscious psychological states and their subsequent actions. The feeling of consciously willing our action, in contrast, is not a direct readout of such scientifically verifiable will power. Rather, it is the result of a mental system whereby each of us *estimates* moment-to-moment the role that our minds play in our actions.

1.2. Mind perception

Why would people mistake the experience of will for a causal mechanism? Why is it that the phenomenal will so easily overrides any amount of preaching by scientists about the mechanisms underlying human action? Now as a rule, when people find an intuition so wildly intriguing that they regularly stand by it and forsake lots of information that is technically more correct, they do so because the intuition *fits.* It is somehow part of a bigger scheme of things that they simply cannot discard. So, for example, people once held tight to the Ptolemaic idea that the sun revolves around the earth, in part because this notion fit their larger religious conception of the central place of the earth in God's universe. In exactly this way, conscious will fits a larger conception – our understanding of *causal agents.*

1.2.1. Causal agency. Most adult humans have a very well-developed idea of a particular sort of entity, an entity that *does things.* We appreciate that a dog, for example, will often do things that are guided not by standard causal principles, but rather by a teleological or purposive system. Dogs often seem to be goal-oriented, as they behave in ways that only seem to be understandable in terms of goals (including some fairly goofy ones, yes, but goals nonetheless). They move toward things that they subsequently seem to have wanted (because they consume them or sniff them), and they move away from things that we can imagine they might not like (because the things are scary or loud or seem to be waving a rolled-up newspaper). Dogs, like horses and fish and crickets and even some plants, seem to be understandable through a special kind of thinking about goal-oriented entities that does not help us at all in thinking about bricks, buttons, or other inanimate objects.

The property of goal seeking is not just something we attribute to living things, as we may appreciate this feature in computers or robots or even thermostats. But the important characteristic of such goal-seeking entities is that we understand them in terms of *where we think they are headed* rather than in terms of *where we think they have been.* Unlike a mere object, which moves or "acts" only when it has been caused to do so by some prior event, a causal agent moves or acts apparently on its own, in the pursuit of some future state – the achievement of a goal. Fritz Heider (1958; Heider & Simmel 1944) observed that people perceive persons as causal agents – origins of events – and that this is the primary way in which persons are understood in a manner that physical objects and events are not.

Causal agency, in sum, is an important way in which people understand action, particularly human action. In the process of understanding actions performed by oneself or by another, the person will appreciate information about intentions, beliefs, desires, and plans, and will use this information in discerning just what the agent is doing. The intuitive appeal of the idea of conscious will can be traced in part to the embedding of the experience of will, and of the notion that will has a force, in the larger conception of causal agency. Humans appear to be goal-seeking agents who have the special ability to envision their goals consciously in advance of action. The experience of conscious will feels like being a causal agent.

1.2.2. Mechanisms and minds. We all know a lot about agents and goals and desires and intentions, and use these concepts all the time. These concepts are only useful, however, for understanding a limited range of our experience. The movements of clock hands and raindrops and electric trains, for example, can be understood in terms of causal relations that have no consciousness or will at all. They are mechanisms. Extending the notion of causal agency to these items – to say these things have the ability to *cause themselves to behave* – does not fit very well with the physical causal relations we perceive all around us. Imagine for a moment a spoon, knife, and fork deciding to go for a walk

https://doi.org/10.1017/S0140525X04000159 Published online by Cambridge University Press

to the far end of the dinner table ("we're off to see the salad . . ."), and you can see the problem. Things do not usually will themselves to move, whereas people seem to do this all the time.

This rudimentary observation suggests that people have at hand two radically different systems of explanation, one for minds and one for everything else. Mentalistic explanation works wonders for understanding minds, but it does not work elsewhere – unless we want to start thinking that everything from people to rocks to beer cans to the whole universe actually does what it consciously wants. Mechanistic explanation, in turn, is just splendid for understanding those rocks and beer cans, not to mention the movements of the planets, but meanwhile leaves much to be wanted in understanding minds.

Each of us is quite comfortable with using these two very different ways of thinking about and explaining events – a physical, mechanical way and a psychological, mental way. In the mechanical explanatory system, people apply intuitive versions of physics to questions of causality, and so they think about causes and effects as events in the world. In the mental explanatory system, in turn, people apply implicit psychological theories to questions of causality, focusing on issues of conscious thoughts and the experience of will as they try to explain actions. In the mechanical way of thinking, all the psychological trappings are unnecessary; a physical system such as a clock, for example, does not have to intend to keep time or to experience doing so. The essence of the mental explanatory system, in contrast, is the occurrence of the relevant thoughts and feelings about the action. In this system, the objects and events of physical causality are not particularly important; a person might experience having willed the death of an enemy and become wracked with guilt, for example, even though there was no mechanism for this to have happened.

These two explanatory systems fall into place as children develop ways of understanding both the physical and psychological worlds. The first inklings that mind perception and mechanistic explanation might develop separately in children came from Jean Piaget, whose perspective has culminated in the contemporary literature on the development of "theory of mind" in animals (Premack & Woodruff 1978) and in children (e.g., Wellman 1992), and in work that contrasts how children develop an understanding of agency, intention, and will with how they develop an understanding of causality, motion, and the principles of physics (e.g., Carey 1996; Gelman et al. 1995). Neither the perception of the physical world nor the perception of the mental world is a "given" to the new human. Although the neonate has rudimentary abilities in both areas, both systems must be developed over time and experience as ways of understanding what all is going on.

The idea that mind perception is variable has also been noted by Dennett (1987; 1996), who captured this observation in suggesting that people take an "intentional stance" in perceiving minds that they do not take in perceiving most of the physical world. The degree to which we perceive mindedness in phenomena can change, such that under some circumstances we see our pet pooch as fully conscious and masterfully deciding just where it would be good to scratch himself, whereas under other circumstances we may have difficulty extending the luxury of presumed conscious thought and human agency even to ourselves. It is probably the case, too, that the degree of mechanical causality we perceive is something that varies over time and circumstance. Viewing any particular event as mentally or mechanically caused, therefore, can depend on a host of factors and can influence dramatically how we go about making sense of it. And making sense of our own minds as mentally causal systems – conscious agents – includes accepting our feelings of conscious will as authentic.

1.3. Real and apparent mental causation

Any magician will tell you the key to creating a successful illusion: The illusionist must make a marvelous, apparently magical event into the easiest and most immediate way to explain what are really mundane events. Kelley (1980) described this in his analysis of the underpinnings of magic in the perception of causality. He observed that stage magic involves a *perceived causal sequence* – the set of events that appears to have happened – and a *real causal sequence* – the set of events the magician has orchestrated behind the scenes. The perceived sequence is what makes the trick. Laws of nature are broken willy-nilly as people are sawed in half, birds and handkerchiefs and rabbits and canes and what-have-you appear from nothing, and also disappear, or for that matter turn into each other and then back again.

The real sequence is often more complicated or unexpected than the illusion, but many of the real events are not perceived. The magician needs special pockets, props, and equipment, and develops wiles to misdirect audience attention from the real sequence. In the end, the audience observes something that seems to be simple, but in fact it may have been achieved with substantial effort, preparation, practice, and thought on the magician's part. The lovely assistant in a gossamer gown apparently floating effortlessly on her back during the levitation illusion is in fact being held up by a 600-pound pneumatic lift hidden behind the specially rigged curtain. It is the very simplicity of the illusory sequence, the shorthand summary that circumvents all the poor magician's toil, which makes the trick so compelling. The lady levitates. The illusion of conscious will occurs by much the same technique (Wegner 2003a).

The real causal sequence underlying human behavior involves a massively complicated set of mechanisms. Everything that psychology studies can come into play to predict and explain even the most innocuous wink of an eye, not to mention some of the more lengthy and elaborate behaviors of which humans are capable. Each of our actions is really the culmination of an intricate set of physical and mental processes, including psychological mechanisms that correspond to the traditional concept of will – in that they involve linkages between our thoughts and our actions. This is the empirical will. However, we do not see this. Instead, we readily accept the far easier explanation of our behavior that our Houdini-esque minds present to us: We think we did it.

Science fiction writer Arthur C. Clarke (1973, p. 21) remarked that "Any sufficiently advanced technology is indistinguishable from magic." Clarke meant this to refer to the fantastic inventions we might discover in the future, or might find if we were to travel to advanced civilizations. However, the insight also applies to self-perception. When we turn our attention to our own minds, we find that we are suddenly faced with trying to understand an unimaginably advanced technology. We cannot possibly know (let alone keep track of) the tremendous number of mechanistic influences on our behavior, because we have the fortune of

inhabiting some extraordinarily complicated machines. So we develop a shorthand – a belief in the causal efficacy of our conscious thoughts. We believe in the magic of our own causal agency.

The mind creates this continuous illusion because it *really doesn't know* what causes its actions. Whatever empirical will there is rumbling along in the engine room – an actual relation between thought and action – might in fact be totally inscrutable to the conscious mind. The mind has a self-explanation mechanism that produces a roughly continuous sense that what is in consciousness is the cause of action – the phenomenal will – whereas in fact the mind actually cannot ever know itself well enough to be able to say what the causes of its actions are. To quote Spinoza in *The Ethics:* "Men are mistaken in thinking themselves free; their opinion is made up of consciousness of their own actions, and ignorance of the causes by which they are determined. Their idea of freedom, therefore, is simply their ignorance of any cause for their actions" (Spinoza 1677/1883, Part II, p. 105). In the more contemporary phrasing of Minsky (1985, p. 306), "none of us enjoys the thought that what we do depends on processes we do not know; we prefer to attribute our choices to *volition*, *will*, or *self-control*. . . . Perhaps it would be more honest to say, *'My decision was determined by internal forces I do not understand'"* (emphasis in original).

2. Apparent mental causation (Ch. 3)

Imagine for a moment that by some magical process, you could always know when a particular tree branch would move in the wind. Just before it moved, you knew it was going to move, in which direction, and just how it would do it. Not only would you know this, but let us assume that the same magic would guarantee that you would happen to be thinking about the branch just before each move. You would look over, and then just as you realized it was going to move, it would do it! In this imaginary situation, you could eventually come to think that you were somehow causing the movement. You would seem to be the source of the distant branch's action, the agent that wills it to move. The feeling that one is moving the tree branch surfaces in the same way that one would get the sense of performing any action at a distance. All it seems to take is the appropriate foreknowledge of the action. Indeed, with proper foreknowledge it is difficult *not* to conclude one has done the act, and the feeling of doing may well-up in direct proportion to the perception that relevant ideas had entered one's mind before the action. This is beginning to sound like a theory.

2.1. A theory of apparent mental causation

The experience of will may be a result of the same mental processes that people use in the perception of causality more generally. The theory of apparent mental causation, then, is this: *people experience conscious will when they interpret their own thought as the cause of their action* (Wegner & Wheatley 1999). This means that people experience conscious will quite independent of any actual causal connection between their thoughts and actions. Reductions in the impression that there is a link between thought and action may explain why people get a sense of involuntariness even for actions that are voluntary, for example, during motor automatisms such as table-turning, or in hypnosis, or in psychologically disordered states such as dissociation. And inflated perceptions of the link between thought and action, in turn, may explain why people experience an illusion of conscious will at all.

The person experiencing will, in this view, is in the same position as someone perceiving causation as one billiard ball strikes another. As we learned from Hume, causation in bowling, billiards, and other games is inferred from the constant conjunction of ball movements. It makes sense, then, that will – an experience of one's own causal influence – is inferred from the conjunction of events that lead to action. Now, in the case of billiard balls, the players in the causal analysis are quite simple: one ball and the other ball. One rolls into the other and a causal event occurs. What are the items that seem to click together in our minds to yield the perception of will?

One view of this was provided by Ziehen (1899), who suggested that thinking of self before action yields the sense of agency. He proposed that

> we finally come to regard the ego-idea as the cause of our actions because of its very frequent appearance in the series of ideas preceding each action. It is almost always represented several times among the ideas preceding the final movement. But the idea of the relation of causality is an empirical element that always appears when two successive ideas are very closely associated. (Ziehen 1899, p. 296)

And indeed, there is evidence that self-attention is associated with perceived causation of action. People in an experiment by Duval and Wicklund (1973) were asked to make attributions for hypothetical events (a hypothetical item: "Imagine you are rushing down a narrow hotel hallway and bump into a housekeeper who is backing out of a room"). When asked to decide who was responsible for such events, they assigned more causality to themselves if they were making the judgments while they were self-conscious. Self-consciousness was manipulated in this study by having the participants sit facing a mirror, but other contrivances – such as showing people their own video image or having them hear their tape-recorded voice – also enhance causal attribution to self (Gibbons 1990).

This tendency to perceive oneself as causal when thinking about oneself is a global version of the more specific process that appears to underlie apparent mental causation. The specific process is the perception of a causal link not only between self and action, but between one's own thought and action. We tend to see ourselves as the authors of an act when we have experienced relevant thoughts about the act at an appropriate interval in advance, and so can infer that our own mental processes have set the act in motion. Actions we perform that are not presaged in our minds, in turn, would appear not to be caused by our minds. The intentions we have to act may or may not *be* causes, but this does not matter, as it is only critical that we *perceive* them as causes if we are to experience conscious will.

In this analysis, the experience of will is not a direct readout of some psychological force that causes action from inside the head. Rather, will is experienced as a result of an interpretation of the *apparent* link between the conscious thoughts that appear in association with action and the nature of the observed action. *Will is experienced as the result of self-perceived apparent mental causation.* Thus, in line with facets of several existing theories (Brown 1989; Clax-

https://doi.org/10.1017/S0140525X04000159 Published online by Cambridge University Press

ton 1999; Harnad 1982; Hoffmann 1986; Kirsch & Lynn 1999b; Langer 1975; Libet 1985; Spanos 1986b; Spence 1996), this theory suggests that the will is a conscious experience that is derived from interpreting one's action as willed. Also in line with these theories, the present framework suggests that the experience of will may only map rather weakly, or at times not at all, onto the actual causal relationship between the person's cognition and action. The new idea introduced here is the possibility that the experience of acting develops when the person infers that his or her own *thought* was the cause of the action.

This theory makes sense as a way of seeing the will because the causal analysis of anything, not only the link from thought to action, suffers from a fundamental uncertainty. Although we may be fairly well convinced that A causes B, for instance, there is always the possibility that the regularity in their association is the result of some third variable, C, which causes both A and B. Drawing on the work of Hume, Jackson (1998) reminds us that "anything can fail to cause anything. No matter how often B follows A, and no matter how initially obvious the causality of the connection seems, the hypothesis that A causes B can be overturned by an over-arching theory which shows the two as distinct effects of a common underlying causal process" (p. 203). Although day always precedes night, for example, it is a mistake to say that day *causes* night, because of course both are caused in this sequence by the rotation of the earth in the presence of the sun.

This uncertainty in causal inference means that no matter how much we are convinced that our thoughts cause our actions, it is still true that both thought and action could be caused by something else that remains unobserved, leaving us to draw an incorrect causal conclusion. As Searle (1983) has put it:

> It is always possible that something else might actually be causing the bodily movement we think the experience [of acting] is causing. It is always possible that I might think I am raising my arm when in fact some other cause is raising it. So there is nothing in the experience of acting that actually guarantees that it is causally effective. (p. 130)

We can never be sure that our thoughts cause our actions, as there could always be causes of which we are unaware, but that have produced both the thoughts and the actions.

This theory of apparent mental causation depends on the idea that consciousness does not know how conscious mental processes work. When you multiply 3 times 6 in your head, for example, the answer just pops into mind without any indication of how you did that. As Nisbett and Wilson (1977) have observed, the occurrence of a mental process does not guarantee the individual any special knowledge of the mechanism of this process. Instead, the person seeking self-insight must employ a priori causal theories to account for his or her own psychological operations. The conscious will may thus arise from the person's theory designed to account for the regular relationship between thought and action (Wegner 2003b). Conscious will is not a direct perception of that relationship, but rather a feeling based on the causal inference one makes about the data that do become available to consciousness: the thought and the observed act.

2.2. Principles of causal inference

How do we go about drawing the inference that our thought has caused our action? Several ideas about this pop up on considering the tree branch example once more. Think, for instance, of what could spoil the feeling that you had moved the branch. If the magic limb moved before you thought of it moving, there would be nothing out of the ordinary and you would experience no sense of willful action. The thought of movement would be interpretable as a memory or even a perception of what had happened. If you thought of the tree limb moving and then something quite different moved (say, a nearby chicken dropped to its knees), again there would be no experience of will. The thought would be irrelevant to what had happened, and you would see no causal connection. And if you thought of the tree limb moving but noticed that something other than your thoughts had moved it (say, a squirrel), no will would be sensed. There would simply be the perception of an external causal event. These observations point to three key sources of the experience of conscious will: the *priority*, *consistency*, and *exclusivity* of the thought about the action (Wegner & Wheatley 1999). For the perception of apparent mental causation, the thought should occur before the action, be consistent with the action, and not be accompanied by other potential causes.

Studies of how people perceive external physical events (Michotte 1963) indicate that the perception of causality is highly dependent on these features of the relationship between the potential cause and potential effect. The candidate for the role of cause must come first or at least at the same time as the effect, it must yield movement that is consistent with its own movement, and it must be unaccompanied by rival causal events. The absence of any of these conditions tends to undermine the perception that causation has occurred. Similar principles have been derived for the perception of causality for social and everyday events (Einhorn & Hogarth 1986; Gilbert 1997; Kelley 1972; McClure 1998), and have also emerged from analyses of how people and other organisms respond to patterns of stimulus contingency when they learn (Alloy & Tabachnik 1984; Young 1995). The application of these principles to the experience of conscious will can explain phenomena of volition across a number of areas of psychology.

2.3. Intentions as previews

The experience of will is the way our minds portray their operations to us, not their actual operation. Because we have thoughts of what we will do, we can develop causal theories relating those thoughts to our actions on the basis of priority, consistency, and exclusivity. We come to think of these prior thoughts as intentions, and we develop the sense that the intentions have causal force even though they are actually just previews of what we may do. Yet, in an important sense, it must be the case that *something* in our minds plays a causal role in making our actions occur. That something is, in the theory of apparent mental causation, a set of unconscious mental processes that cause the action. At the same time, that "something" is very much like the thoughts we have prior to the action.

One possibility here is that thought and action arise from coupled unconscious mental systems. Brown (1989) has suggested that consciousness of an action and the performance of the action are manifestations of the same "deep structure." In the same sense that the thought of being angry might reflect the same underlying process as the experience of facial flushing, the thought and performance of a

https://doi.org/10.1017/S0140525X04000159 Published online by Cambridge University Press

voluntary action might be different expressions of a singular underlying system. The coupling of thought and action over time in the adult human is really quite remarkable if the thought is *not* causing the action, so there must be some way in which the two are in fact often connected.

The co-occurrence of thought and action may happen because thoughts are normally thrust into mind as *previews* of what will be done. The ability to know what one will do, and particularly to communicate this to others verbally, would seem to be an important human asset, something that promotes far more effective social interaction than might be the case if we all had no idea of what to expect of ourselves or of anyone around us. The thoughts we find coming to our minds in frequent coordination with what we do may thus be produced by a special system whose job it is to provide us with ongoing verbalizable previews of action. This preview function could be fundamentally important for the facilitation of social interaction. Intentions, in this analysis, are to action what turn signals are to the movements of motor vehicles. They do not cause the movements, they preview them.

By this logic, real causal mechanisms underlying behavior are never present in consciousness. Rather, the engines of causation operate without revealing themselves to us, and so may be unconscious mechanisms of mind. The research suggesting a fundamental role for automatic processes in everyday behavior (Bargh 1997) can be understood in this light. The real causes of human action are unconscious, so it is not surprising that behavior could often arise – as in automaticity experiments – without the person having conscious insight into its causation. Conscious will itself arises from a set of processes that are not the same processes as those that cause the behavior to which the experience of will pertains, however. So, even processes that are not automatic – mental processes described as "controlled" (Posner & Snyder 1975) or "conscious" (Wegner & Bargh 1998) – have no direct expression in a person's experience of will. Such "controlled" processes may be less efficient than automatic processes and require more cognitive resources, but even if they occur along with an experience of control or conscious will, this experience is not a direct indication of their real causal influence. The experience of conscious will is just more likely to accompany inefficient processes than efficient ones because there is more time available prior to action for inefficient thoughts to become conscious, thus to prompt the formation of causal inferences linking thought and action. This might explain why controlled/conscious processes are often linked with feelings of will, whereas automatic processes are not. Controlled and conscious processes are simply those that lumber along so inefficiently that there is plenty of time for previews of their associated actions to come to mind and allow us to infer the operation of conscious will (Wegner 2005).

The unique human convenience of conscious thoughts that preview our actions gives us the privilege of feeling we willfully cause what we do. In fact, however, unconscious and inscrutable mechanisms create both conscious thought about action and the action as well, and also produce the sense of will we experience by perceiving the thought as cause of action. So, although our thoughts may have deep, important, and unconscious causal connections to our actions, the experience of conscious will arises from a process that interprets these connections, not from the connections themselves.

3. The mind's compass (Ch. 9)

Does the compass steer the ship? In some sense, you could say that it does, because the pilot makes reference to the compass in determining whether adjustments should be made to the ship's course. If it looks as though the ship is headed west into the rocky shore, a calamity can be avoided with a turn north into the harbor. But of course, the compass does not steer the ship in any physical sense. The needle is just gliding around in the compass housing, doing no actual steering at all. It is thus tempting to relegate the little magnetic pointer to the class of epiphenomena – things that do not really matter in determining where the ship will go.

Conscious will is the mind's compass. As we have seen, the experience of consciously willing action occurs as the result of an interpretive system, a course-sensing mechanism that examines the relations between our thoughts and actions and responds with "I willed this" when the two correspond appropriately. This experience thus serves as a kind of compass, alerting the conscious mind when actions occur that are likely to be the result of one's own agency. The experience of will is therefore an indicator, one of those gauges on the control panel to which we refer as we steer. Like a compass reading, the feeling of doing tells us something about the operation of the ship beneath us. But also like a compass reading, this information must be understood as a conscious experience, a candidate for the dreaded "epiphenomenon" label. Just as compass readings do not steer the boat, conscious experiences of will do not cause human actions.

Why is it that the conscious experience of will exists at all? Why, if this experience is not a sensation of the personal causation of action, would we even go to the trouble of having it? What good is an epiphenomenon? The answer to this question becomes apparent when we appreciate conscious will as a feeling that organizes and informs our understanding of our own agency. Conscious will is a signal with many of the qualities of an emotion, one that reverberates through the mind and body to indicate when we sense having authored an action. The idea that conscious will is an *emotion of authorship* moves beyond the standard way in which people have been thinking about free will and determinism and presses toward a useful new perspective.

3.1. Free will and determinism

A book called *The Illusion of Conscious Will* certainly gives the impression of being a poke in the eye for readers who believe in free will. It is perfectly reasonable to look at the title and think the book is all about determinism and that it will give the idea of free will no fair hearing at all. And, of course, the line of thought here does take a decidedly deterministic approach. For all this, though, our discussion has actually been *about* the experience of free will, examining at length when people feel it and when they do not. The special idea we have been exploring is to explain the experience of free will in terms of deterministic or mechanistic processes.

On the surface, this idea seems not to offer much in the way of a solution for the classic question of free will and determinism. How does explaining the feeling of will in terms of deterministic principles help us to decide which one is true? Most philosophers and people on the street see this

https://doi.org/10.1017/S0140525X04000159 Published online by Cambridge University Press

as a fight between two big ideas, and they call for a decision on which one is the winner. As it turns out, however, a decision is not really called for at all. The usual choice we are offered between these extremes is not really a choice, but rather a false dichotomy. It is like asking: Shall we dance, or shall we move about the room in time with the music? The dichotomy melts when we explain one pole of the dimension in terms of the other. Still, this does not sit well with anyone who is married to the standard version of the problem, so we need to examine just how this usual choice leads us astray.

3.1.1. The usual choice. Most of us think we understand the basic issue of free will and determinism. The question seems to be whether all our actions are determined by mechanisms beyond our control, or whether at least some of them are determined by our free choice. Described this way, many people are happy to side with one possibility or the other. There are those of us who side with free will, and thus view members of the opposition as nothing but *robogeeks,* creatures who are somehow disposed to cast away the very essence of their humanity and embrace a personal identity as automatons. There are others of us, however, who opt for the deterministic stance, and thus view the opposition as little more than *bad scientists,* a cabal of confused mystics with no ability to understand how humanity fits into the grand scheme of things in the universe. Viewed in each others' eyes, everyone comes out a loser.

The argument between these two points of view usually takes a simple form: The robogeeks point to the array of evidence that human behavior follows mechanistic principles, taking great pride in whatever data or experiences accumulate to indicate that humans are predictable by the rules of science. Meanwhile, the bad scientists ignore all of this and simply explain that their own personal experience carries the day. They know they have conscious will. And no one wins the argument. The usual clash fails on both sides because free will is a feeling, whereas determinism is a process. They are incommensurable. Free will is apples and determinism is oranges.

The illogic of treating free will and determinism as equal opposites becomes particularly trenchant when we try to make free will do determinism's causal job. What if, for example, we assume that free will is just like determinism, in that it is also a process whereby human behavior can be explained? Rather than all the various mechanistic engines that psychologists have invented or surmised in humans that might cause their behavior, imagine instead a person in whom there is installed a small unit called the Free Willer. This is not the usual psychological motor, the bundle of thoughts or motives or emotions or neurons or genes – instead, it is a black box that just *does things.* Many kinds of human abilities and tendencies can be modeled in artificially intelligent systems, after all, and it seems on principle that we should be able to design at least the rudiments of a psychological process that has the property of freely willing actions.

But what exactly do we install? If we put in a module that creates actions out of any sort of past experiences or memories, that fashions choices from habits or attitudes or inherited tendencies, we do not get freedom, we get determinism. The Free Willer must be a mechanism that is *unresponsive to any past influence.* In *Elbow Room: The Varieties of Free Will Worth Wanting,* Dennett (1984) illustrates how hollow and unsatisfying free will of this kind might be. In essence, any such system makes sense only if it inserts some fickle indeterminacy into the person's actions. Dennett points out that it is not particularly interesting or fun to have a coin flipper added to the works somewhere between "sensory input" and "behavior output." Who would want free will if it is nothing more than an internal coin flip? This is not what we mean when we talk about our own conscious will. Trying to understand free will as though it were a kind of psychological causal process leads only to a mechanism that has no relationship at all to the experience of free will that we each have every day.

People appreciate free will as a kind of personal power, an ability to do what they want to do. Voltaire (1694–1778) expressed this intuition in saying, "Liberty then is only and can be only the power to do what one will" (1752/1924, p. 145). He argued that this feeling of freedom is not served at all by the imposition of randomness, asking, "would you have everything at the pleasure of a million blind caprices?" (p. 144). The experience of will comes from having our actions follow our wishes, not from being able to do things that do not follow from anything. And, of course, we do not cause our wishes. The things we want to do come into our heads. Again quoting Voltaire, "Now you receive all your ideas; therefore you receive your wish, you wish therefore necessarily. . . . The will, therefore, is not a faculty that one can call free. The free will is an expression absolutely devoid of sense, and what the scholastics have called will of indifference, that is to say willing without cause, is a chimera unworthy of being combated" (p. 143). A Free Willer, in short, would not generate the experience of conscious will.

We are left, then, with a major void. In leaving out a mechanism that might act like free will, theories have also largely ignored the experience of free will. The feeling of doing is a profoundly regular and important human experience, however, and one that each of us gets enough times in a day to convince us that we are doing things (non-randomly) much of the time. This deep intuitive feeling of conscious will is something that no amount of philosophical argument or research about psychological mechanisms can possibly dispel. Even though this experience is not an adequate theory of behavior causation, it needs to be acknowledged as an important characteristic of what it is like to be human. People feel will, and scientific psychology needs to know why. Clearly, people do not feel will because they somehow immediately know their own causal influence as it happens. The experience is the endpoint of the very elaborate inference system underlying apparent mental causation, and the question becomes: Why do we have this feeling?

3.1.2. Authorship emotion. Perhaps we have conscious will because it helps us to appreciate and remember what we are doing. The experience of will marks our actions for us. It helps us to know the difference between a light we have turned on at the switch and a light that has flickered alive without our influence. To label events as our personal actions, conscious will must be an experience that is similar to an emotion. It is a feeling of doing. Unlike a cold thought or rational calculation of the mind alone, will somehow happens both in body and in mind. The experience of willing an action has an embodied quality, a kind of weight or bottom, which does not come with thoughts in general. In the same sense that laughter reminds us that our bodies are

having fun, or that trembling alerts us that our bodies are afraid, the experience of will reminds us that we are doing something. Will, then, serves to accentuate and anchor an action in the body. This makes the action our own far more intensely than could a thought alone. Unlike simply saying "this act is mine," the occurrence of conscious will brands the act deeply, associating the act with self through feeling, and so renders the act one's own in a personal and memorable way. Will is a kind of authorship emotion.

The idea that volition is an emotion is not new. In fact, T. H. Huxley (1910) made the equation explicit: "Volition . . . is an emotion *indicative* of physical changes, not a *cause* of such changes. . . . The soul stands to the body as the bell of a clock to the works, and consciousness answers to the sound which the bell gives out when struck. . . . We are conscious automata." Will is a feeling, not unlike happiness or sadness or anger or anxiety or disgust. Admittedly, conscious will does not have a standard facial expression associated with it, as do most other basic emotions. The look of determination or a set brow that is sometimes used to caricature willfulness is probably not identifiable enough to qualify as a truly communicative gesture. Still, will has other characteristics of emotion, including an experiential component (how it feels), a cognitive component (what it means and the thoughts it brings to mind), and a physiological component (how the body responds). Although conscious will is not a classic emotion that people would immediately nominate when asked to think of an emotion, it has much in common with the emotions.

The experience of consciously willing an action belongs to the class of *cognitive feelings* described by Gerald Clore (1992). He points out that there is a set of experiences such as the feeling of knowing, the feeling of familiarity, or even the feeling of confusion, that serve as indicators of mental processes or states, and that thus inform us about the status of our own mental systems. The experience of willing an action is likewise an informative feeling, a perception of a state of the mind and body that has a unique character. Although the proper experiments have not yet been done to test this, it seems likely that people could discriminate the feeling of doing from other feelings, knowing by the sheer quality of the experience just what has happened. The experience of willing is more than a perception of something outside oneself, it is an experience of one's own mind and body in action.

Conscious will is the emotion of authorship, a somatic marker (Damasio 1994) that authenticates the action's owner as the self. With the feeling of doing an act, we get a conscious sensation of will attached to the action. Often, this marker is quite correct. In many cases, we have intentions that preview our action, and we draw causal inferences linking our thoughts and actions in ways that track quite well our own psychological processes. Our experiences of will, in other words, often do correspond correctly with the empirical will, the actual causal connection between our thought and action. The experience of will then serves to mark in the moment and in memory the actions that have been singled out in this way. We know them as ours, as authored by us, because we have felt ourselves doing them. This helps us to tell the difference between things we are doing and all the other things that are happening in and around us. In the melee of actions that occur in daily life, and in the social interaction of self with others, this body-based signature is a highly useful tool. We resonate with what we do, whereas we only notice what otherwise happens or what others have done – so we can keep track of our own contributions without pencils or tally sheets.

Conscious will is particularly useful, then, as a guide to our selves. It tells us what events around us seem to be attributable to our authorship. This allows us to develop a sense of who we are and are not. It also allows us to set aside our achievements from the things that we cannot do. And perhaps most important for the sake of the operation of society, the sense of conscious will also allows us to maintain the sense of responsibility for our actions that serves as a basis for morality.

We can feel moral emotions inappropriately, of course, because our experience of conscious will in any given case may be wrong. The guilt we feel for breaking mother's back may accrue via the nonsensical theory that we were culpable for her injury as a result of stepping on a crack. More realistically, we can develop guilty feelings about all sorts of harms we merely imagine before they occur – simply because our apparent mental causation detector can be fooled by our wishes and guesses into concluding that we consciously willed events that only through serendipity have followed our thoughts about them. By the same token, the pride we feel in helping the poor may come from the notion that we had a compassionate thought about them before making our food donation, although we actually were just trying to clear out the old cans in the cupboard. But however we do calculate our complicity in moral actions, we then experience the emotional consequences and build up views of ourselves as certain kinds of moral individuals as a result. We come to think we are good or bad on the basis of our authorship emotion. Ultimately, our experience of conscious will may have more influence on our moral lives than does the actual truth of our behavior causation.

3.2. How things seem

Sometimes how things seem is more important than what they are. This is true in theater, in art, in used car sales, in economics, and, it now turns out, in the scientific analysis of conscious will as well. The fact is, it seems to each of us that we have conscious will. It seems we have selves. It seems we have minds. It seems we are agents. It seems we cause what we do. Although it is sobering and ultimately accurate to call all this an illusion, it is a mistake to conclude that the illusory is trivial. To the contrary, the illusions piled atop apparent mental causation are the building blocks of human psychology and social life. It is only with the feeling of conscious will that we can begin to solve the problems of knowing who we are as individuals, of discerning what we can and cannot do, and of judging ourselves morally right or wrong for what we have done.

Usually, we assume that how things seem is how they are. We experience willing a walk in the park, winding a clock, or smiling at someone, and the feeling keeps our notion of ourselves as persons intact. Our sense of being a conscious agent who does things comes at a cost of being technically wrong all the time. The feeling of doing is how it seems, not what it is – but that is as it should be.

3.3. Postscript

This précis of *The Illusion of Conscious Will* is an abridgement of three of the book's chapters. It focuses on the main

https://doi.org/10.1017/S0140525X04000159 Published online by Cambridge University Press

arguments, and leaves aside the bulk of the empirical evidence relevant to these arguments. The evidence is essential and extensive, however, and the arguments cannot be evaluated effectively without it. Like a vertebrate stripped of its skeleton, this article does not stand on its own.

To prop up the arguments here, or at least to see where they might stand if they were ossified, several lines of evidence can be noted. One key theme of the book is the analysis of automatisms – actions experienced as occurring without conscious will. A variety of historical examples of automatisms from the Spiritualist literature of the nineteenth century (e.g., automatic writing, pendulum divining, Ouija board spelling), along with more contemporary research on the role of automaticity in everyday action (e.g., Bargh & Ferguson 2000), reveal the frequent occurrence of voluntary action without experienced conscious will. The case of hypnosis is also examined in depth, as a means of establishing some of the conditions under which people lose conscious will while still performing complicated, goal-directed actions.

The flip side of such under-experience of will is, of course, the *over*-experience of will – the feeling of will for actions the person did not perform. Evidence for such erroneously inflated will is found in the psychological literature on perceived control and the illusion of control (Haidt & Rodin 1999; Langer 1975; Taylor & Brown 1988). There is also evidence indicating that the over-experience of will occurs as predicted by the principles of the theory of apparent mental causation (Ansfield & Wegner 1996; Wegner & Wheatley 1999).

Another line of evidence on conscious will involves the construction of agents. When people fail to experience will even while performing complicated voluntary actions, they often attribute the performance to other agents (although these agents could not have performed the action). The book examines the creation of such *virtual agency* in a number of domains, reviewing evidence on the attribution of actions to both real and imaginary agents. When people in 1904 became convinced that the horse Clever Hans was accurately answering their questions with his hoof tapping, for example – whereas in fact the horse was responding to their unconscious nonverbal communication of the answers – they were projecting their own actions on another agent. The related case of facilitated communication, in which people helping others to communicate fail to appreciate their own contribution to the communication, also illustrates the extraordinary mutability of the experience of will (see also Wegner et al. 2003). The lack of conscious will in such unusual phenomena as spirit possession and dissociative identity disorder is explored, too, as these cases also involve the construction of virtual agents as the person's way of understanding actions not consciously willed by the agent self.

A final body of evidence on illusory will has to do with the cognitive distortions that operate to protect the illusion. Studies of the confabulation of intention following action show that people often invent or distort thoughts of action in order to conform to their conception of ideal agency. People who are led to do odd actions through post-hypnotic suggestion, for example, often confabulate reasons for their action. Such invention of intentions is the basis for a variety of empirical demonstrations associated with theories of cognitive dissonance (Festinger 1957) and the left-brain interpretation of action (Gazzaniga 1983). Operating on the assumption that they are agents leads people to presume that they intended actions even when this could not have been the case, to misperceive their actions as being consistent with their intentions, and to experience conscious will whenever their intentions and actions happen to coincide.

The idea that conscious will is an illusion, in sum, is supported by a range of experimental and case demonstrations of the extraordinary dissociation of the experience of will and the actual wellsprings of action. People feel will for actions they did not cause, and can feel no will for actions they clearly did cause. The fundamental disconnection of the feeling from the doing suggests that the feeling of conscious will issues from mental mechanisms that are not the same as the mental mechanisms that cause action.

ACKNOWLEDGMENTS

The full set of acknowledgments for this work appears in *The Illusion of Conscious Will.* Here it must suffice to thank Thalia Wheatley for her contribution to the theory, to thank Heather Gray for comments on this target article, and to note that the preparation of the article was supported by NIMH Grant 49127.

Open Peer Commentary

The self is virtual, the will is not illusory

George Ainslie

Veterans Affairs Medical Center, Coatesville, PA 19320.
george.ainslie@med.va.gov **www.picoeconomics.com**

Abstract: Wegner makes an excellent case that our sense of ownership of our actions depends on multiple factors, to such an extent that it could be called virtual or even illusory. However, two other core functions of will are initiation of movement and maintenance of resolution, which depend on our accurate monitoring of them. This book shows that will is not an imponderable black box but, rather, an increasingly accessible set of specific functions.

This book is an encyclopedic analysis of the ways in which our sense of volition fools us. Wegner (2002) has assembled a remarkably broad range of examples wherein people behave without being aware of deciding to do so; falsely believe that they are deciding; or, most subtly, experience a decision as occurring at a different time than objective evidence places the decision. I think that Wegner over-reads the implications of these examples when he calls conscious will an illusion. Our eyes sometimes fool us, too, as when we mislocate an underwater object or are led by contextual cues to misjudge the size or distance of an object, but we still say that we are actually seeing it. The famous moon illusion does not make the moon illusory. Wegner has many valuable things to say, but the examples he assembles to argue against conscious will apply to only parts of what his own material demonstrates to be a complex phenomenon. I submit that what he – and we – call "conscious will" comprises at least three somewhat independent processes, two of which depend on the person's accurate sense of their operation.

Dealing with these two first: The initiation of movement and the maintenance of resolution, perhaps Wegner's "little dabs" of will and its "long lasting property," respectively, each has its kind of proprioception within the mind (brain?) itself; we rely on the accuracy of this proprioception from minute to minute, day in and