

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 automatism – actions experienced as occurring without conscious will. A variety of historical examples of automatism 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

day out. One of Wegner's own examples illustrates the *initiation of movement* part of the will to move. The amputee who is conscious of moving nonexistent toes is obviously not relying on peripheral sensations. She reports mentally doing what, in someone with toes, accurately governs their movements. By abnormally removing the peripheral component of this process, nature has isolated Hume's "*impression we feel and are conscious of, when we knowingly give rise to any new motion of our body*" (Hume 1739/1888, p. 399, quoted in sect. 1.1.1 of the target article, emphasis in original). The associated movements are gone, but the experience of will in this trivial sense of connecting mind and body remains, and there is no reason to believe that the subject's consciousness of its operation *per se* is inaccurate, despite the illusory downstream effects. This consciousness is different in kind from mere association; if a tree branch actually moved without my proprioception of will every time I thought of its moving, it would not feel as if I suddenly had a previously unrecognized muscle, but instead would probably give me the eerie sensation of having my mind read (see Gray Walter's experiment in Dennett 2003a, p. 240).

Maintenance of resolution is more important. It is where both strength and freedom of will reside, and our beliefs about it have practical effects on self- and social control. Defending direct perception of this resolution is hard because, although observers have agreed on many functional properties – the effects of practice, of reference to principles, of single lapses, and so forth (Ainslie 2001, pp. 119–20) – they have not agreed on a way of describing the thing itself. I have argued that resolution is not a thing, or unitary sensation, at all, but an intertemporal process analogous to bargaining, and that it is just as directly reportable as the events of interpersonal bargaining are (Ainslie 2001, pp. 90–104). Briefly: The way we make our intentions consistent is to perceive our current decision as a test case for how we will decide similar choices generally, so that our expected reward from consistent intention is staked on "cooperating" with our future selves and is sharply reduced if we "defect" to an impulsive alternative. Although people conceive the mechanics of this contingency variously, under the rubrics of morality, principle, personal intention, and even divine help, we universally experience a big stake as resolve and a lapse as a loss of part of this stake, engendering guilt. The proprioception here is the recursive self-monitoring process, the testing of our will, which is not prominent in behaviors we are confident of executing but is glaringly evident when we resolve to resist a favorite vice or to dive into a cold lake. The mind's compass to which Wegner refers (sect. 3) is not the same thing as our will but, rather, is a component of it, as integral as the thermometer is to a thermostat. Furthermore, the sensitive dependence of our behavior on our compass readings – the feedback prospective outcome of tentative choices – is enough to account for the experience of freedom, our sense that we are participating in the outcome but that even we cannot be sure of its final form in advance.

Is there an illusion, then? A penetrating chapter on "virtual agency" (not in the *Précis*) suggests a more defensible illusion, involving a *third part* of the experience of will – neither the part that connects mind to action in little dabs nor the long-lasting property that manages resolve, but the part that connects our actions with our idea of our selves. The evidence of this chapter indicates that it is not our sense of action that is illusory (I like "virtual" better), but, rather, our sense of self. Wegner argues for possibilities that I have also advocated: that a person interprets her own actions in the same way she interprets others' – *empathically*, as I put it – so that the ownership of both kinds of action and the notion of ownership itself are open to construction, and facts without major practical implications are chosen for belief on the basis of how regularly they occasion emotion (Ainslie 1995; 2001, pp. 175–89). Wegner says that the conscious will departs when people feel possessed or depersonalized; that they have lost their empathic sense of self, their "emotion of authorship," leading them to feel that they do not own their activities. Nevertheless, these people continue to perform consciously the other two functions of will: initiating actions and maintaining resolutions. The ownership compo-

nent could indeed be called illusory or virtual or emotional, but it is not essential for the functioning of conscious will.

Most of the examples of failed consciousness in the book depend on either a split of consciousness or activity below a threshold of consciousness. The splits remove the reporting self's "emotion" of agency by physically (split brain; alien hand) or motivationally (dissociation and probably hypnosis) blocking this partial self's awareness of what are often fully formed initiations and resolutions. Subthreshold phenomena include mannerisms (which can be shaped even in sleep, Granda & Hammack 1961); small drifts of activity that can be summed into Ouija-like phenomena; and the preliminary brain processes made tangible by recent advances in neurophysiology and imaging. We can now see a decision in its early stages, perhaps when it is merely being mooted and not yet a decision – the "mirror neurons" excited by watching somebody else's movements do not always, or even usually, result in your own actual movement (Iacoboni et al. 1999); perhaps Libet's electrodes (1999) are also registering the first idea of a behavior and not the decision to go forward with it, a possibility that would reduce the significance of the observed temporal offset from the conscious moment of choice. With powerful cranial magnets we can even skip the perceptual phase of suggestion and predispose directly to one alternative over another (Brasil-Neto et al. 1992), but the capacity to manipulate an early stage of will does not argue against its existence. Science sees submerged parts of an iceberg that have never been seen before, but as yet nothing that renders the conscious parts inaccurate.

The wealth of material in this book – brain imaging, electrophysiology, social experiments, anthropological observations, and thought experiments – demonstrates that the will is not a unitary organ with no discernable components and an either/or outcome structure, the black box traditionally favored by philosophers (e.g., Pap 1961, p. 213). It is divisible into separate operations, some of which can be measured as lasting finite, very short times. These elements may relate to one another in a variety of ways, including, as I have suggested, in recursive feedback systems, while being experienced only as summation phenomena – an experience that is incomplete, as Wegner demonstrates, but normally valid as far as it goes. What used to be called conation turns out to be a field as big as cognition. This book goes a long way toward defining its tasks.

NOTE

The author of this commentary is employed by a government agency and as such this commentary is considered a work of the U.S. government and not subject to copyright within the United States.

The experience of will: Affective or cognitive?

Joseph E. Bogen

Neurologic Surgery, University of Southern California, Los Angeles, CA 90033, and Division of Biology, California Institute of Technology, Pasadena, CA 91105.
jbogen@its.caltech.edu <http://www.its.caltech.edu/~jbogen>

Abstract: Wegner vacillates between considering the experience of will as a directly-sensed feeling and as a cognitive construct. Most of his book is devoted to examples of erroneous cognition. The brain basis of will as an immediately-sensed emotion receives minimal attention.

Wegner sometimes considers the experience of will to be "a feeling" (Wegner 2002, p. 3), directly sensed, "not unlike happiness or sadness" (p. 326). However, he more often considers it a "fabrication" (p. 3), a cognitive construct combining what he calls priority, consistency, and exclusivity (p. 69). For some of us, the idea that will is a directly-sensed feeling suggests a search for the neural correlates of this feeling (Bogen 1997). Wegner briefly refers to stimulation of the exposed cortex by Penfield and the well-confirmed result that the movements elicited are disclaimed by the patients. He contrasts this with a single case of Delgado wherein