

Wegner provides us with case after case of how our sensations of will are mistaken, how we sometimes do things ourselves but attribute these actions to others, how we sometimes think we are doing things ourselves, but we aren't. He is right; our sense of will is sometimes – maybe a lot of the time – misleading.

But so what? What, if anything, does this tell us about freedom of the will – the actual will, not what we sense as a marker for the will? Not much. In order to know something about the actual will, for example, whether it exists in any interesting sense, we would have to know how the sensation of will connects up with either our underlying psychology or our underlying physiology or both. However, unlike the case of hunger, in which we know a lot about the connection between various levels of hormones in our blood stream and wanting to eat, we know very little about what the sensation of will actually reflects. Maybe it does mirror a genuine self in the brain: our central control that initiates or at least approves our purposeful behaviors. Maybe it doesn't. But knowing that our conscious sensations of will are sometimes mistaken doesn't shed any light on this topic.

We know some actions happen to us – I sneezed in the middle of lecture – and others have a psychological reason behind them – I raised my hand in the middle of lecture. We can tell the difference between these sorts of activities, both from the inside, as it were, and from the outside. But what is this difference? Is it just that the latter is accompanied by a sensation of will and the former isn't? Is it just that we explain the latter in terms of beliefs and desires and the former in terms of physiology? Or does the latter occur as the endpoint in a causal chain mediated by my own psychological states, whereas the former doesn't? I think that no matter what one's metaphysical stripe, one would have to agree with the last suggestion: What differentiates willed actions from actions that are not willed is the causal history of the action. Willed actions flow from or through my psychological states in ways that unwilled actions don't.

But if this is the case, then in what sense is our sensation of conscious will an illusion? Our sensation serves to differentiate which actions flowed from or through our psychological economy from those that did not. It may get it wrong once in a while; it may get it wrong lots of times. Nevertheless, the sensation is reflecting something real, as real as our bodies' need for nutrients. The important question is what exactly is that sensation reflecting.

Wegner wants to argue that we don't really have selves, that our sensation of selfhood, too, is just another inference our bodies and brains make about what we are doing in order to explain our selves to ourselves. He wants to argue that we have this whole edifice of illusory constructions about our own psychologies from which the sensation of will flows. He wants to argue these things, but he can't. He can't because he doesn't get below the sensations to learn what is really going on. He has "Just So" stories about how selves might work, but so do a lot of people. Unless and until we can connect our sensations to actual physiological or deeper psychological workings, it will be hard to claim that our sensation of will is illusory in any interesting sense.

The sense of conscious will

Gene M. Heyman

Behavioral Psychopharmacology Research Laboratory and Psychiatry, McLean Hospital and Harvard Medical School, Belmont, MA 02478.
gmheyman@mclean.harvard.edu

Abstract: Wegner's conclusion that conscious will is an illusion follows from a key omission in his analysis. Although he describes conscious will as an experience, akin to one of the senses, he omits its objective correlate. The degree to which behavior can be influenced by its consequences (voluntariness) provides an objective correlate for conscious will. With conscious will anchored to voluntariness, the illusion disappears.

When an object, say a boat on the water, moves away, its retinal image decreases in size. However, instead of experiencing the boat

as shrinking, the viewer experiences it as receding into the distance. This could be called an illusion; the retinal image is getting smaller, not further away. However, to say that this is an "illusion" is to ignore the determinants of object constancy. When the viewer's understanding of boats and the three-dimensional world are included in the analysis of visual experience, the correlation between a shrinking retinal image and the perception of a constant sized but increasingly distant object is perfectly understandable. Or, to put it another way, to say that one of the visual constancies is an "illusion" is to overlook that there is more to vision than the retina.

Wegner's treatment of conscious will (Wegner 2002) is rather like trying to account for object constancy while limiting the analysis to the retina. He emphasizes that conscious will is an inference and that its contents often do not match up well with the actual factors that cause voluntary action. For instance, we may be aware of the intention to raise our hand (or assume this intention after the fact), but not be aware of the determinants of this intention or of having made an inference. From these "discrepancies," Wegner concludes that conscious will is an illusion. However, as in the object constancy example, a more complete account of the input eliminates the illusion.

Object constancy is about the fact that we live in a three-dimensional world and that when objects move, they usually do not change shape. What is conscious will about? What is its stimulus? The answer cannot be found in a textbook (as with the perceptual constancies), but it is familiar and easily identified.

As documented by Wegner, conscious will's domain is behavior, in particular our own behavior. Just as perception tracks dimensions of the external world, conscious will tracks the important fact that our own activities vary in the degree to which they are influenced by consequences (e.g., rewards, incentives, punishments, and the like), by the values we adopt, and by new information. Some activities are immune to these factors, whereas others are easily modified by just a hint of praise or disapproval. For instance, consider the different causal relations relating to a patellar reflex and learning to kick a ball, blushing and the decision to wear rouge, a defensive blink and a conspiratorial wink at a friend. The second activity in each comparison we call voluntary, and the first we call involuntary. The distinction is not a matter of free will versus determinism. Antecedents govern voluntary and involuntary acts. Rather, the mediating neural architecture and nature of the antecedents differ. Differences in neural connections allow for variation in the degree to which activities are influenced by experience and the contents of consciousness. The distinction also does not depend on intentions or other subjective reports. We can be conscious of involuntary acts (I know I am going to blink, but I can't help it), and as Wegner's literature review ably demonstrates, we can be unaware of voluntary acts. In other words, voluntariness (susceptibility to consequences) provides an objective basis for subjective experience, just as the conservation of an object's shape and size while moving provides a basis for perceptual constancy.

Wegner acknowledges that behavior varies with regard to its susceptibility to consequences (e.g., the ear wiggling discussion, Wegner 2002, pp. 31–34), and also acknowledges that voluntary actions are the usual focus of conscious will. However, these observations are made in passing, and his analysis proceeds without any further discussion of the objective basis for the sensation of "doing something." Given this omission it is understandable that he concludes that it is an illusion. This is not to say that conscious will is a literal reflection of susceptibility to rewards. For instance as Wegner's discussion of automatic processes (2002, pp. 56–59) demonstrates, many learned, reinforced actions can move out of awareness.

Leaving out the objective correlates for conscious will leads to empirical and logical problems. An empirical shortcoming is the de-emphasis of the contribution that conscious will makes to voluntary action. Often Wegner seems to be saying that conscious will is no more than an after-the-fact frill, at best useful for a kind of moral bookkeeping (see below). I am not sure that this is what he

really means, because it is easy to show that the contents of consciousness (e.g., plans) can alter the course of voluntary action. To make a less obvious point: Voluntary behavior is subject to competing contingencies, and without conscious awareness of the more global ones (those that are good for us in the long run), we would always fall victim to the most immediate reward (e.g., Heyman 1996; 2003). This observation is celebrated in Greek myth (e.g., the story of Odysseus and the sirens) and is embedded in moral and spiritual teachings (which can be seen in part as pleas for attending to vital but less salient, long-term contingencies).

The logical problem is that leaving out the stimulus leads to circular accounts of how conscious will arises and its purpose. According to Wegner, associative correlations are sufficient for the sense of conscious will. He writes that if intentions or plans are followed by action, the mind infers that the action was caused by conscious will. However, intentions and plans imply the sensation of conscious will. In order to have an intention or plan regarding action, one must already have the belief that behaviors exist that can be modified by goals and consequences. By leaving out the behavioral basis for plans (voluntariness), his account amounts to the circular statement that conscious will (intentions and plans) is the basis of conscious will.

There is a similar logical problem with Wegner's theory of what conscious will is good for – the purposes it serves. He asks (p. 325): Why do we have the feeling of conscious will? His general conclusion is that it serves as a guide for moral responsibility. His argument is that we should only be responsible for the actions that we intended, and, hence, that conscious will serves the purposes of the moral order. However, the vocabulary of moral interactions assumes “intentions” and “free choices.” The view of personal responsibility that Wegner appeals to assumes the existence of conscious will, and more fundamentally, it assumes the existence of activities that are modifiable by intentions and consequences (such as the fear of punishment). Hence, moral responsibility (as defined by Wegner) cannot explain why we have the feeling of conscious will; it builds on its prior existence. The way out of this circularity is to identify the objective correlates of the sensation of will.

Wegner's literature review and his own experiments make it clear that conscious will is in many respects like the basic senses. Like the senses it is correlated with an important dimension of the objective world; like the senses it does not provide a literal representation of either the objective world or the proximal stimulus; and like the senses it has proven a useful guide to more effective voluntary actions. My title for the experimental literature Wegner reviews would have been: *The sense of conscious will*.

ACKNOWLEDGMENT

I thank Martha Pott for her helpful comments.

How neuroscience accounts for the illusion of conscious will

Masao Ito

Brain Science Institute, Institution RIKEN, Wako, Saitama, 351-0198, Japan.
masao@brain.riken.jp

Abstract: Wegner's monograph presents the view that conscious will is a feeling that we experience when we perform an action through a mechanistic process of the brain, rather than a mental force that causes the action. The view is supported by several lines of evidence in which conscious will is dissociated from the actual performance of voluntary movements, as in automatism. The book further extends an insightful analysis of the mental system behind the illusion of conscious will and inspires neuroscientists to reflect on its neural substrates.

Wegner's (2002) book challenges a core issue in the brain-mind problem, that is, *conscious will*. It will be a milestone in our extensive effort to clarify the brain-mind relationship. Chapter One

clarifies two contrasting views of the conscious will. The first view is that the conscious will is a mental force that we traditionally believe to be causal to a voluntary action. The second view is that the conscious will is a feeling associated with a mechanistic brain process that causes a voluntary action. The book presents a comparison of the two views, and inclines to discard the first view as an illusion. The succeeding chapters collect various observations of human brains and mentations to substantiate the arguments. Several lines of reported examples such as the alien hand, hypnotic experiences, acts due to spiritualism, and the phantom limb indicate that conscious will can be dissociated from actual voluntary movement, contrary to the first view. With the first view discarded, how can we explain the contradictory situation where we feel we are willfully causing an action that is in fact a product of a certain mechanistic unconscious brain process? The book analyzes various possible sources of the illusion of conscious will. A frequent coincidence between an intention and the actual action can be mistaken as implying causality; prediction of an action before it happens may lead to a feeling that our will is causing the action; and an intention can be confabulated after an action has been performed.

Explanations from the internal model hypothesis. This book inspires neuroscience to find neuronal counterparts of the seemingly mysterious mental processes reported. I find such a neuronal counterpart in the internal model hypothesis proposed in cerebellar neuroscience. In brief, the hypothesis assumes that the cerebellum forms an internal model, which, by subtle learning mechanisms of cerebellar neuronal circuits, copies functional properties of a motor apparatus that the motor cortex controls (Ito 1984). The internal model provides an internal feedback to the motor cortex, and thereby enables us to perform a movement even with an impaired sensory feedback (Wegner 2002, p. 39). The cerebellar internal model may also assist in predicting sensory consequences of movement (Miall et al. 1993). A movement may accompany sensation, which disturbs the performance of the movement. The cerebellar internal model may predict and subsequently block such a disturbance. This hypothesis explains our experience that a self-generated tactile stimulus is perceived as less ticklish than a similar stimulus applied externally (Blakemore et al. 2000). A phantom limb (Wegner 2002, p. 40) may likewise occur if there be a mismatch between external sensory and cerebellum-mediated internal feedbacks. If a cerebellar internal model reproduces inverse dynamics properties of a motor apparatus, it can replace the controller function of the motor cortex (Kawato et al. 1987). Because the cerebellum is not involved in brain mechanisms underlying consciousness, in which the cerebral cortex is generally involved, a learned movement can then be performed unconsciously through the cerebellar pathway. This condition closely resembles that illustrated in Figure 3.1 (Wegner 2002, p. 68) for explaining conscious and unconscious mental events behind voluntary movement.

The internal model hypothesis has been expanded to problems of the thought (Ito 1993). When we think, the prefrontal cortex acts as an executive cortex and manipulates an image, a concept, or an idea, which are collectively termed the mental model. A mental model is formed by combining various pieces of information received from the sensory cortex, and is stored in the temporoparietal cortex that constitutes the internal environment of the brain. Just as we manipulate an arm or a leg during movements, we manipulate a mental model during thoughts. During repeated trials of thought, a mental model in the temporoparietal cortex is copied in a cerebellar internal model. By referring to such a copy of a mental model, the thought can be performed quickly and unconsciously of its processes. This fits the situation described on page 67: “when you multiply 3 times 6 in your head, the answer just pops into your mind without any indication of how you did that” (Wegner 2002). Such a calculation, when first performed in the cerebral cortex, must require conscious effort, but as a learnt calculation it is performed in the cerebellum, and will no longer rise to the level of consciousness.