

Is conscious will an illusion?*

Jing Zhu

Chinese Academy of Sciences

Abstract

In this essay I critically examine Daniel Wegner's account of conscious will as an illusion developed in his book *The Illusion of Conscious Will* (MIT Press, 2002). I show that there are unwarranted leaps in his argument, which considerably decrease the empirical plausibility and theoretical adequacy of his account. Moreover, some features essential to our experience of willing, which are related to our general understanding of free will, moral responsibility and human agency, are largely left out in Wegner's account of conscious will. This substantially diminishes its implications and significance for some profound philosophical issues.

1. Introduction

Normally, we have the experience that we consciously will what we do: we are the authors or originators who make our actions happen. Actions are generally understood as what we voluntarily do, perform or bring about, instead of what we undergo or merely happen to us. We do things, and when we do them, we experience the actions as coming

* I would like to thank Andrei Buckareff for many helpful comments on the first draft of this essay. Earlier versions of this paper were presented at *The 95th Annual Meeting of the Southern Society for Philosophy and Psychology*, on April 17, 2003, in Atlanta, Georgia, *2003' APA Central Division Meeting*, on April 25, 2003, in Cleveland, Ohio, and *The 29th Annual Meeting of the Society for Philosophy and Psychology*, on June 19, 2003, in Pasadena, California. I am grateful to the audiences for valuable feedback, and especially, to Eddy Nahmias and Thomas Polger, who were my commentators respectively at the latter two meetings.

seamlessly from our thoughts (e.g., intentions) and consciousness. The experience of willing indicates a sense of agency that some of our conscious thoughts causally bring about our actions.

In a rich and fascinating book, *The Illusion of Conscious Will*, the psychologist Daniel Wegner argues that conscious will, which encompasses both the experience of willing and the perception of the causation of action by conscious thought, is in fact an illusion. 'It is an illusion in the sense that the experience of consciously willing an action is not a direct indication that the conscious thought has caused the action' (p. 2). This idea is not entirely new. David Hume, for example, conceives 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' (*Treatise of Human Nature*, II, 3, 1, 2). Well informed by historical and contemporary psychological studies and neuroscientific research, Wegner develops a detailed and engaging account on how this illusion is created by the mind and brain. According to Wegner, the illusion of conscious will arises when people mix up the experience of willing and the apparent causal efficacy of our conscious thoughts on our actions. So conscious will is essentially a construction or fabrication. But it is a useful illusion, for it signals personal authorship of action to the individual and therefore serves as the basis for personal achievement and moral responsibility.

In this essay, I will critically examine Wegner's account of conscious will as an illusion. I will show that there are unwarranted leaps in Wegner's argument, which considerably decrease the empirical plausibility and theoretical adequacy of his account. Moreover, some distinctive features of the experience of willing, which are essential to our general understanding of free will, moral responsibility and human agency, are largely left out in Wegner's discussion of conscious will. This substantially diminishes its implications and significance for some profound philosophical issues.

2. The generation of conscious will

How does conscious will arise? Here is a sketch of the story that Wegner tells in great detail in his book: First, the agent has a certain thought, most commonly an intention. Then the agent experiences a voluntary action. If there is a constant conjunction between the occur-

rence of the thought and that of the action, the agent tends to perceive that the thought caused the action, just like his perceiving causation in general as one billiard ball strikes another. ‘People experience conscious will when they interpret their own thought as the cause of their action’ (Wegner 2002, p. 64). That is,

When we think that our conscious intention has caused the voluntary action that we find ourselves doing, we feel a sense of will. We have willfully done the act. (p. 68)

Some specifications are in order. First, for an agent to infer that his own thought causes his action, the relation between the thought and the action must satisfy some requirements. Three principles on the relation between the thought and the action are essential to support the perception of mental causation: the *priority*, *consistency*, and *exclusivity* of the thought about the action (Wegner and Wheatley 1999; Wegner 2002, ch. 3). More specifically, the thought should occur before the action, be consistent with the action, and not be accompanied by other potential causes. Similar principles have been derived for the perception of causality for external physical and social events. And the absence of any of these conditions tends to undermine the perception that causation has occurred. ‘The application of these principles to the experience of conscious will can explain phenomena of volition across a number of areas of psychology’ (Wegner 2002, p. 70).

Secondly, Wegner thinks that the experience of willing is necessary for an agent to act voluntarily (2002, pp. 3-4):

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 isn’t willed if the person says it was not. . . . Consciously willing an action requires a feeling of doing, a kind of internal ‘oomph’ that somehow certifies authentically that one has done the action.

In short, the experience of willing, which may consist of a feeling of doing and a sense of authorship of one’s own action, is essential for an agent to perceive some of his bodily movements as voluntary actions.

Wegner’s account of the generation of conscious will, however, is insufficient in several respects. First, our experience of willing does not

exclusively depend on the perception of causal efficiency of our thought on our action. Sometimes even if the intended action fails to occur in accordance with our intention, we can still have a sense of willing. For instance, in an attempt to move a heavy box, your strength may not be strong enough to make the intended movement occur, but you can still experience a feeling of doing or trying. In *The Principles of Psychology*, William James described a case:

Close the patient's eyes, hold his anaesthetic arm still, and tell him to raise his hand to his head; and when he opens his eyes he will be astonished to find that the movement has not taken place. (1981, p. 1101)

Why was this patient surprised? Because he tried to raise his hand, and under normal condition, his hand should be over his head; but unknown to him, his paralyzed arm failed to move in accordance with his intention and attempt. He tried, but failed. But even if he cannot physically raise his arm, he can still try to move it. It seems that a typical voluntary action consists of at least two distinctive components: an inner mental event or process of *trying*, and the corresponding overt bodily movement. By viewing actions in this way, it seems that we can provide a suitable answer to the well-known question put forward by Wittgenstein: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm? (1953, sec. 621) The answer is trying. Trying seems to be essential for an agent to experience that he willed his action.

Wegner suggests that the experience of will could be a result of the same mental processes that people use in the perception of causality more generally: 'The person experiencing will, in this view, is in the same position as someone perceiving causation as one billiard ball strikes another' (2002, p. 64). The similar principles of priority, consistency, and exclusivity are also embodied in the perception of causality in general. However, whereas the causal perception of the movements of billiard balls will not generate any sense of agency—the experience of one's own *doing* something—why is the constant conjunction between one's own thought and action, from which apparent mental causation is inferred, *sufficient* to generate the experience of conscious will and the sense of agency? That is, through the same mental processes that people use in the perception of causality in general, one can infer apparent mental causation (e.g., my intention to raise my arm causes my arm's raising) from the

constant conjunction between a thought and an action, but how can the experience of willing and the sense of agency, which are totally absent in the perception of causality in general, be generated by the *same* mental processes? It seems more plausible to assume that some additional mental processes may be essential in generating the experience of willing.

In many intentional actions, the contribution of the causing intentions is not limited to triggering or initiating the intended actions. Intentions can also play roles of sustaining, guiding and controlling in intentional actions (Adams and Mele 1989; Mele 1992, ch. 8). For example, in writing a book, you cannot simply start the task and then let the process operate by itself; you have to make a continuous voluntary effort to keep going with the activity to its completion. Thus your intention to complete a book not only causally initiates but also help to sustain, guide and control your actions. The ballistic conception of mental causation, according to which the causal contribution of the suitable mental antecedents does not extend beyond triggering an intentional action, appears not suitable to capture the *interactive, continuous* causal relationship between intentions and intentional actions, therefore is insufficient in accounting for the sources of conscious will.

In sum, Wegner's account of the generation of conscious will falls short as a satisfactory explanation for how our experience of willing arises, because it ignores some mental activities that are essential in generating the experience of willing and the sense of doing.

3. Why is conscious will illusory?

Wegner rightly observes that the causal analysis of anything, not only on the link from thought to action, suffers from a fundamental uncertainty. He says: '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' (2002, p. 66). According to his theory of apparent mental causation, both conscious thought about action (e.g., intention, belief) and voluntary action are caused by unconscious mental processes, and the unconscious mental processes which give rise to the conscious thought about an action and which actually cause the voluntary action are probably the same. There is regularity between the thought and the action, but they are both caused by a third factor, namely, some unconscious mental processes.

When we (mistakenly) think that our conscious thought has caused the voluntary action, we experience conscious will.

There are two major sources from which Wegner draws support for his theory of apparent mental causation. One is the study of automatic mental processes. Automatic mental processing occurs reflexively whenever certain triggering conditions are in place—when those conditions are met, the process runs autonomously until its completion, independent of intentional initiation and conscious guidance. Perhaps the most well-known phenomenon of automaticity in human cognition is the Stroop effect, first described by J. R. Stroop in 1935 (Stroop 1935). This is one of the most robust effects in experimental psychology (reviewed by MacLeod 1991). If you try to name the physical color of a word, the semantic processing of the word can ‘interfere’ with the process of naming the color of the word. If the meaning of the word matches its color (e.g., ‘RED’ written in red colour), it takes little time to respond. If the word and color do not match (e.g., ‘RED’ written in green colour), it takes a much longer time to name the color and the accuracy drops dramatically. The most commonly accepted explanation of the Stroop effect is the *automatic word recognition hypothesis*: exposure to a word automatically elicits the unconscious operation that processes the semantic content of the word, which interferes with the more effortful color recognition task. Reading is one of the basic skills that we have well acquired and over-practiced. For a skilled reader, reading a word becomes an unconscious, automatic process, which is hard to be suppressed intentionally.

Recent findings in psychology show that automatic processes penetrate in almost every aspect of our mental life, and play a fundamental role in producing and controlling our cognition, thought, emotion, feeling, and behavior (Bargh 1997, Bargh and Chartrand 1999, Bargh and Ferguson 2000; Wegner and Bargh 1998, Wegner 2002, ch. 4). This suggests to Wegner that ‘if conscious will is illusory, automatism is somehow the ‘real thing,’ fundamental mechanisms of mind that are left over once the illusion has been stripped away. Rather than conscious will being the rule and automatism the exception, the opposite may be true: Automatism is the rule, and the illusion of conscious will is the exception.’ (Wegner 2002, p. 143) However, not all automatic processes are unconscious, unintentional and uncontrollable (cf. Bargh 1989). For example, as Logan and Cowan have pointed out, many processes that are

widely considered to be automatic, such as typing, reading, driving and walking, are actually highly controlled, in that they can be intentional and stoppable (1984). Furthermore, even if a great deal, even the vast majority of human thinking, feeling, and behavior operate in automatic fashion with little or no need for conscious, intentional control, the evidence does not guarantee that the conscious will is an illusion or epiphenomenon. The conscious, controlled mental processes may account for only a small part of the total amount of our mental life, like the tip of iceberg, but this part is nonetheless significant. Conscious, controlled activities are essential in such tasks as planning, decision-making, error-correction or 'trouble-shooting,' learning and acquiring new skills, and overriding strong habitual responses or resisting temptations (Posner and Snyder 1975; Norman and Shallice 1986). The ability of conscious control of one's thought and action can provide substantive adaptive advantages that help a species deal with the environmental and social pressures (Dennett 2003, ch. 8).

Another empirical source that Wegner relies on heavily is the neurophysiologist Benjamin Libet's experimental study on the neural mechanisms and phenomenology of voluntary, self-initiated action. In a series of well-known experiments, Libet found that voluntary actions are initiated by unconscious brain activities well before intentions or decisions to act are consciously experienced by the subjects (1985). In Libet's words, our voluntary actions are 'initiated by unconscious cerebral processes before conscious intention appears' (1985, p. 529). Thus it is our unconscious brain processes, rather than our conscious decisions or intentions, which originally bring about voluntary actions. It seems that both our experience of conscious intentions and the corresponding voluntary actions are caused by specific unconscious brain processes. Therefore, it is wrong to infer that it is the conscious intention that causes the action. The interpretation of Libet's experiments, however, is very controversial. Elsewhere, I have argued that Libet's interpretation of his experiments, in spite of its wide acceptance, is fatally flawed, and there is a much more plausible alternative which can preserve the classical image that we are the originators of our own actions (Zhu 2003; see also Dennett 2003, ch. 8 for an incisive criticism of Libet's account). The cause of the intentional acts performed by the subjects in Libet's experiment should not be *exclusively* attributed to special cerebral processes, like what Libet does; the conscious intention formed at the beginning of experiment when the subjects

received experimental instructions, which guide the subjects' purposeful, effortful and intentional behavior, must be taken into account.

The two major empirical studies that Wegner uses to support his theory of apparent mental causation, which in turn renders crucial support to his account of conscious will as an illusion, are thus either indirect and indecisive, or problematic.

4. Conscious will unexamined

If conscious will is, as Wegner tries to show us, an illusion based on mistaken causal inference about the cause of action, is it useless or even misleading? Wegner contends that it need not be a mere epiphenomenon. On the contrary, it is a useful illusion.

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. ... Just as compass readings do not steer the boat, conscious experiences of will do not cause human actions. (Wegner 2002, p. 317)

According to Wegner, 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' (p. 325). Conscious will is thus particularly useful as a guide to our social life. Although conscious will does not indicate the actual causal process of action, it can enhance our positive experience of personal effectiveness and achievement, which is good for our mental health in many ways. The intentions and conscious thoughts about our actions are cues to ourselves and to others about the likely occurrence of our behavior. The identification of ownership of action helps people attribute responsibility and develop various moral sentiments.

However, if we take Wegner's compass analogy seriously, we will find that the function of conscious will and its implications may be much more significant than what he describes. The compass does not steer the ship by

itself. But it can become a crucial component of a control system, which consists of the pilot and other detecting and effecting devices, causally determining the ship's course. Whether there is a compass, and whether it functions normally, can substantially affect the ship's behavior. Likewise, conscious experiences of will do not cause human actions, but they can become part of a control system that causally determines an agent's behavior, in which conscious will can causally contribute to the control of action. So conscious will not only serves as an indication of authorship of action, it also plays a substantive, causally functional role in the generation and control of our actions, which provides the basis for a full-fledged notion of responsibility. The intentions and conscious thoughts about our actions are not merely cues for ourselves and others to predict our behavior, they are also some of the causal factors that shape and control our actions.

In his recent book *Rationality in Action*, John Searle points out that there are (at least) three gaps in human practical reasoning and action (Searle 2001, pp. 14-15):

First, there is the gap of rational decision making, where you try to make up your mind what you are going to do. Here the gap is between the reasons for making up your mind, and the actual decision that you make. Second, there is a gap between the decision and the action. Just as the reasons for the decision were not causally sufficient to produce the decision, so the decision is not causally sufficient to produce the action. There comes the point, after you have made up your mind, when you actually have to do it. And once again, you cannot sit back and let the decision cause the action, any more than you can sit back and let the reasons cause the decision. ... There is a third gap that arises for actions and activities extended in time, a gap between the initiation of the action and its continuation to completion. ... Even once you have started you cannot let the causes operate by themselves; you have to make a continuous voluntary effort to keep going with the action or activity to its completion.

For an agent, the relation between his thoughts and actions may not be that a conscious thought about action *causes* the intended action, but rather that the agent does something to bridge the gaps between the thought and the action, making the action occur to implement the thought. The role of the agent is thus not merely a passive observer or bystander who just perceives or infers the causal relation between his own thoughts and actions, but rather an active doer or agent *par excellence*

who brings about or tries to bring about the actions in accordance with his intentions. Therefore, the experience of willing or conscious will is likely built on the mental activity by which an agent bridges the gaps between his thoughts and actions, instead of the mere causal perception or inference about the relation between his own thoughts and actions. What is the mental act or activity by which an agent bridges the gaps between his thought and action? It is traditionally conceived as 'volition', 'act of will', or called 'trying' by some contemporary philosophers (see Ginet 1990, ch. 2; Hornsby 1980; Lowe 1995, ch. 5; McCann 1974; Zhu forthcoming a, forthcoming b). A volitional theory of action casts an agent in his proper role, whereas Wegner's account treats a person as a bystander, who passively watches and experiences psychological and physiological events occurring inside him, without taking any active part.

An agent's activity to bridge the gaps between thought and action may not be the only source from which conscious will arises. An equally significant place to look for the origin of the experience of willing is the first gap in practical reasoning that Searle describes, which is between reasons for decision and the decision or the intention to be made or formed. As Searle puts it, you cannot simply sit back and wait for the reasons, which consists of your beliefs and desires, to cause the decision or the intention. You must *do* something to make up your own mind. This is the place where the freedom of will is traditionally supposed to locate. The freedom of will implies not only the freedom to do what one has already decided to do, but also the freedom to make the decision. The experience of the freedom of conscious will implies the sense that an agent feels that it is up to him to make the choice among alternatives of courses of action. One of the main goals of Wegner's book is 'to explain the experience of free will in terms of deterministic or mechanistic processes' (p. 318). However, an essential part of our experience of free will is largely left out in his account of conscious will.

Is conscious will an illusion? It may be, if it is indeed a result of the interpretative processing about the (simplistic) causal relation between our thoughts and actions, as Wegner has shown in *The Illusion of Conscious Will*. It may not be, if our conscious will functions as part of the whole action control system that generates a much more robust sense of agency and ownership, and responds literally with 'I *willed* this'. This latter sense of conscious will remains largely unexamined in Wegner's work.

Jing Zhu (humanwill@yahoo.ca)
 Faculty of Humanities and Social Sciences
 Graduate School of the Chinese Academy of Sciences
 19 (A) Yuquan Road
 Beijing, 100039, China

References

- Adams, F. and A. R. Mele. 1989: The role of intention in intentional action. *Canadian Journal of Philosophy*, 19, 511-32.
- Bargh, J. A. 1989: Conditional automaticity: Varieties of automatic influence in social perception and cognition'. In J. S. Uleman and J. A. Bargh (eds.), *Unintended Thought*. New York: Guilford, 3-51.
- Bargh, J. A. 1997: The automaticity of everyday life. In R. S. Wyer (ed.), *Advances in Social Cognition*. Mahwah, NJ: Lawrence Erlbaum Associates, Vol. 10: 1-61.
- Bargh, J. A. and T. L. Chartrand 1999: The unbearable automaticity of being. *American Psychologist*, 54, 462-479.
- Bargh, J. A. and M. J. Ferguson 2000: Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, 126, 925-945.
- Dennett, D. C. 2003: *Freedom Evolves*. London: Allen Lane.
- Ginet, C. 1990: *On Action*. Cambridge: Cambridge University Press.
- Hornsby, J. 1980: *Actions*. London: Routledge & Kegan Paul.
- Libet, B. 1985: Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Science*, 8, 529-566.
- Logan, G. D. and W. B. Cowan 1984: On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review*, 91, 295-327.
- Lowe, E. J. 1996: *Subjects of Experience*. Cambridge: Cambridge University Press.
- MacLeod, C. M. 1991: Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 193-203.
- McCann, H. 1974: Volition and basic action. *Philosophical Review*, 451-473.
- Mele, A. R. 1992: *Springs of Action: Understanding Intentional Behavior*. Oxford: Oxford University Press.
- Norman, D. A. and T. Shallice 1986: Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz and D. Shapiro (eds.), *Consciousness and Self-Regulation: Advances in Research and Theory*. New York: Plenum, Vol. 4: 1-18.
- Posner, M. I. and C. R. Snyder 1975: Attention and cognitive control. In R. L. Solso (ed.), *Information Processing and Cognition*. Hillsdale, NJ: Erlbaum, 55-85.
- Searle, J. 2001: *Rationality in Action*. Cambridge, MA: MIT Press.
- Stroop, J. R. 1935: Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 28, 643-662.

- Wegner, D. M. 2002: *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Wegner, D. M. and J. A. Bargh 1998: Control and automaticity in social life. In D. T. Gilbert, S. T. Fiske and G. Lindzey (eds.), *Handbook of Social Psychology*. Boston: McGraw-Hill.
- Wegner, D. M. and T. Wheatley 1999: Apparent mental causation: Sources of the experience of will. *American Psychologist*, 54, 480-492.
- Zhu, J. 2003: Reclaiming volition: An alternative interpretation of Libet's experiment. *Journal of Consciousness Studies*, Vol. 10, Issue 11, 61-77.
- Zhu, J. forthcoming a: Understanding volition. *Philosophical Psychology*.
- Zhu, J. forthcoming b: Locating volition. *Consciousness and Cognition*.