

No One is Tricking Anyone: A Critique of Wegner's Theory of Conscious Will

Eelke Spaak (eelke.spaa@gmail.com)
Radboud University Nijmegen

Abstract

Daniel Wegner has demonstrated that the experience of consciously willing an action and the actual causal history of that action sometimes come apart. From this, he concludes that the experience of conscious will and the force of conscious will are entirely distinct and that conscious will is an illusion. This conclusion is, I argue, unjustified, because (1) concluding that something is caused by something else, not just concerning volition and action, is always an inference, and (2) there is no such thing as 'direct perception'. An alternative account is developed in which persons are considered the actual causes of their actions. Wegner's claim that the experience of conscious will arises from a process of interpretation is, in a somewhat modified form, accepted as being most likely true.

Introduction¹

'Free' or 'conscious' will is something with which we humans seem very intimately acquainted. Its nature appears to be almost self-evident: whenever we *do* something, we exercise our free will; we, as rational agents, exert control over our own actions and we initiate them as we see fit. However, over the course of history, this *prima facie* intimacy has been called into question by various disciplines. Theology was among the first: if God already knows everything that is going to happen in advance, what free choice is there left for me to make? In modern times, physics plays a more prominent role: if all there is atoms bumping into each other², how can I freely will something to happen if everything is already determined by the laws that govern the bumping? With the advent and rapid development of contemporary cognitive neuroscience, a new variety of this question has emerged: is my behavior controlled by *me*, or is it 'simply' the output of some complex neurobiochemical machinery processing and subsequently reacting to external stimuli?

In *The Illusion of Conscious Will*, a recent and influential take on this problem, Daniel Wegner (2002) claims that the conscious will people appear to have is actually an illusion. To back up this claim, he has reviewed an extensive body of literature that, among other things, shows that the experience

¹This paper was written in July 2008, and my views on some of its subject matter have changed since then. Specifically, were I to write the same paper again, I would substitute the present, to a great extent Humean, account of causation with one that, on the one hand, appeals even more to sensory experience, and, on the other hand, employs a more sophisticated ontology of objects and properties, that allows for causation on and between different levels. I have Baker (2007) to thank for this. The brunt of the presented argument and its conclusion would remain the same, however. (ES, February 2009)

²Or: if all there is is elementary particles interacting with each other according to a small set of fundamental laws of nature, ...

of having willed something and the actual causal history of that same something sometimes come apart. I find there is great merit in his work, in which he demonstrates virtuosity in intertwining theoretical analyses with empirical findings and their interpretations, and, in doing so, provides important new insights into the nature of conscious will.

However, I believe his main conclusion is unjustified for two reasons. First, it seems he does not take his own distinction between will as an experience and will as a causal force seriously, instead degrading the latter to a variety of the former well before investigating what anyone should mean by it. This mistake seems to stem from too narrow a view on causality, and while he often refers to David Hume's famous treatment of the problem of causality (Hume, 1739/1978, 1777/2005), he does not, I think, quite draw the conclusions he should from Hume's analysis. Second, when humans perceive *anything*, they always do this through their sense organs. Wegner sometimes talks about 'direct perception', but, in fact, there is no such thing.

In the present paper, I first give an overview of Wegner's account of conscious will, as presented in his 2002 book, *The Illusion of Conscious Will*, as well as in some subsequent writings (Wegner, 2003, 2005), especially devoting attention to his views on how the experience of conscious will arises and to his theory of 'apparent mental causation'. After this, I will use Wegner's texts and the two objections mentioned above as a basis for developing an alternative view on the nature of conscious will.

Wegner's Theory

Wegner (2002) begins *The Illusion of Conscious Will* by stating that there are two main ways of explaining the behavior of any given human being. First, one could investigate that particular human being's psychological mechanisms, genetic makeup, neuroanatomical structure, history of learning, etc. and then predict his or her next move. A fundamental assumption of psychology is, according to Wegner, that, given a complete description of all such factors, predictions like these will be completely accurate. A rather different way of explaining human behavior is by saying that the human being displaying the behavior has consciously willed that behavior. These two ways of explaining behavior are usually at odds: the mechanistic explanation is the explanation of choice for science, while the 'conscious will' explanation is usually arrived at after intuitive introspection.

Wegner goes on to analyze the 'conscious will' explanation. Conscious will is usually understood in one of two ways: as an

experience or as a causal force. These two aspects of conscious will are, Wegner claims, entirely distinct, and “the tendency to confuse them is the source of the illusion of conscious will” (Wegner, 2002). Wegner presents his views on the causal force of conscious will more or less intertwined with his views on its relation with the experience of conscious will. Because his views on this relation will be treated in the next section (‘An Alternative View’), his views on will as a causal force will be treated in that section, while his comments on conscious will as an experience will be described here.

Will as an Experience

Will is a feeling, and it is this quality that makes first-person reports of the experience of it incorrigible by any external measures. Wegner asks us to consider, for example, the case of alien hand syndrome, a disorder most frequently occurring in patients having undergone a split-brain procedure. People with alien hand syndrome report having no voluntary control over one of their arms (see, e.g., Scepkowski & Cronin-Golomb, 2003, for a recent review), yet external observers report “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” (Wegner, 2002). Wegner claims that the only reason we do not classify the alien hand’s actions as willful is the patients’ reports that the actions lack some sort of internal ‘oomph’ experience that would normally accompany voluntary actions initiated by them.

In the remainder of his book, Wegner reviews a number of other cases to show that the experience of having willed an action need not actually correspond to either the appearance of the action as willful to external observers, or to the actual causal history of the action. His entire review will not be repeated here, but one example, that of *table-turning*, is worth mentioning, because the actual causal history of an action was quantitatively determined and found to differ from the experience of conscious will as reported by participants. Table-turning involves participants sitting around a table with their hands touching the table’s rim and waiting for the table to start moving ‘by itself’. Indeed, the table often starts moving miraculously. Michael Faraday (1853) investigated this phenomenon by installing force-measuring devices between the table and the participants’ hands, and found that the participants, still denying that the movement was caused by them, were actually exerting force on the table.

Because the experience of conscious will and the actual control of a person’s actions sometimes come apart, Wegner proposes to make a distinction between two factors: *action* (‘doing’/‘not doing’) and *feeling of will* (‘feeling of doing’/‘no feeling of doing’). There are four possible combinations of the levels of these factors, two of them resulting in normal, everyday action or inaction, in which it is accompanied by the appropriate feeling of action or inaction, and two of them resulting in abnormal conditions: automatism (doing without feeling of doing) and illusion of control (feeling of doing without actually doing). This is, I believe, a useful distinction and the reality of all four combinations is highly plausible, given the empirical evidence presented by Wegner (2002).

Apparent Mental Causation

From this dissociation between action and feeling of doing, Wegner concludes that, while people usually experience their actions as being caused by them consciously willing those actions, this is in reality not the case. Instead, the illusory experience of conscious will arises when people *interpret* their own thoughts as the causes of their actions. Conscious will thus arises from a process of causal inference (Wegner, 2003).

This causal inference, Wegner concludes from the empirical data on automatisms and illusions of control, happens according to three principles: priority, consistency, and exclusivity. These hold that, for an action to be interpreted by a person as being caused by his or her thoughts, i.e., as being willed by him or her, the thought has to occur just before the action (priority); the thought and action have to be consistent³ (consistency); and there should be no other potential causes of the action present (exclusivity).

Because of this inference, people experience an *apparent* causal link between their thoughts and their actions, while, in fact, both the actions and the thoughts are caused by unconscious mental events, and it is the process of interpreting the covariation of thought and action according to the three principles of priority, consistency, and exclusivity, that gives rise to the illusion of conscious will.

An Alternative View

Wegner’s study nicely demonstrates the fact that the actual causal history of an action and the reported feeling of conscious will by the person executing that action need not always agree. Therefore, I wholeheartedly agree with Wegner in concluding that, indeed, our first-person experience of conscious will is an interpretation of the actual state of affairs in the real world. However, his conclusion from this that conscious will is an illusion is, I think, the wrong one.

As already mentioned in the introduction, I think Wegner does not take his own distinction between will as an experience and will as a causal force seriously. He writes: “it is tempting to think that the conscious experience of will is a direct perception of the force of will” (Wegner, 2002). Wegner rightly argues that this is not the case, but he does not properly investigate the possibility that conscious experience of will is a *not-so* direct perception of the force of will and that, therefore, will as an experience and will as a causal force might not be “entirely distinct”, as he claims they are.

I think the conscious experience of the will should be considered an experience of the causal force of the will, even though this experience can, at times, be imperfect. To understand conscious will, we first need to devote some attention to the nature of causality. Second, we need to know what the causal force of the will should be taken to be, and, finally, we need to understand how the conscious experience of this causal force could have arisen. Each of these points will now be addressed.

³Wegner defines this principle less clearly than the others. To illustrate the point, he writes: “when people do what they think they were going to do, there exists consistency between thought and act” (Wegner, 2002).

Causality

Wegner often refers to David Hume's famous treatment of the problem of causality (Hume, 1739/1978, 1777/2005). He paraphrases Hume as saying, with regard to causality in the interaction among objects, that "causality is not a property inhering in objects" but that it is "an event, not a thing" (Wegner, 2002). In giving Hume's view on mental causation, Wegner writes: "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" (Wegner, 2002). Wegner thus seems to take causality, when inferred from the interaction between two physical objects, as in some ways more 'real' (an "event" of some kind) than the causality that is sometimes inferred from the conjunction between will and action.

Hume himself would certainly disagree. He writes, on comparing willed and involuntary causation: "is it more difficult to conceive that motion may arise from impulse than that it may arise from volition? All we know is our profound ignorance in both cases" (Hume, 1777/2005). I think Hume is right in asserting that causality is *always* an inference; an inference that we are quite good at making and that we often make without realizing it, but an inference nonetheless.

Another quote of Wegner's will make the problem and its solution clearer. When one claims that conscious will is "a force in a person that causes the person's action", one is, according to Wegner, giving "the same kind of explanation as saying that God has caused an event." Such an explanation "is a stopper that trumps any other explanation." (Wegner, 2002). I believe this analogy is fundamentally flawed, since the rationale behind an act of God is, by definition, completely beyond our comprehension, whereas an act that is humanly willed (regardless of whether this should be understood as some sort of 'real' causal process or as 'just' an experience) can be analyzed, both by interrogating the person that has acted ("I did *this* because of *that*"), as well as investigating the circumstances of his decision from a scientific, empirical, viewpoint. These analyses do not preclude the description of the act as willed (again, regardless of will being illusory or not), so the description of the act as willed certainly does not trump any other explanation. The implicit assumption that seems to underlie Wegner's giving this analogy seems to be that, once the cause of an event is known, no other cause can be given for this same event. This, I would say, is only true when one is including some 'final' and 'uncaused' cause (i.e., God) in one's explanation, and not in any other causal description of an event.

When a billiards cue strikes the cue ball, which in turn strikes the eight-ball, certainly the cue ball's bumping into the eight-ball can be considered as having caused the eight-ball's subsequent moving, but does this mean that one is wrong when asserting that the eight-ball's moving has been caused by the cue having stricken the cue ball? It would seem not, since this striking obviously is an essential element in the chain of events leading up to the eight-ball's moving. Of course, one could argue that there are 'causal intermediaries' (i.e., the cue ball's moving about) involved, so the causation is not 'direct', but this is also the case with the cue ball bumping into the eight-ball: the bumping first causes a very rapid elastic 'squeezing' of both billiard balls, which in turn causes an equal expansion, which in turn causes them to move apart. Such

analyses can be taken further, and, in the end, they might arrive at some sort of 'fundamental' causation (most likely involving atoms, their positions, charges, etc.), but we still cannot be sure that even *this* causation resides *in* the atoms or their interactions; it is still an attribution, or inference, that we make.

This is not really a problem, however, since the inference is usually correct. Basic experimental research methodology tells us that, whenever an experimenter manipulates only a single factor (and controls for other factors, uses naive test subjects, etc.) and measures another factor and finds that this second factor lawfully varies with his manipulations of the first factor, he is justified in concluding the experimental manipulations have caused the variation in the measured variable. This is an intuitively plausible formalization of causal inference and I believe there are few who would want to question it.

Infants, who certainly are unaware of any such methodological considerations, are, already at a young age, capable of making causal inferences. For example, when a block is placed on top of another block, and the bottom block is subsequently removed, they will expect the resulting lack of support to cause the top block to fall down (Baillargeon, Kotovsky, & Needham, 1995). While causation is not necessarily and with a deductive certainty present in the real world, it is not surprising that infants, just like all humans, are capable of such a feat: the world certainly *seems* to be governed by causal principles, and since the human species has evolved in this highly-causally-seeming world and each individual grows up in this same world, it is only natural to expect the processes of phylogenetic natural selection and ontogenetic development to exploit this regularity of the world in order to better cope with it.

So, summarizing this section so far, causality is *always* an inference from observations, but, in most cases, humans are, both intuitively⁴ and in more formalized manners⁵, quite good at performing this inference. The question of whether or not we should call causation as we use it in our descriptions of the world a *real* property of nature itself is, I think, entirely beside the point: we just as well could call *all* such causation real, or *all* such causation illusory.

The Will

Having analyzed what we mean by causality, we can investigate the causal force the will might have.

To quote Hume again, one possible definition of a person is as "nothing but a bundle or collection of different perceptions" (Hume, 1739/1978), and while Wegner certainly disagrees with Hume on the "nothing but" part, he does seem to agree with him that the "bundle or collection of different perceptions" is the most important aspect of what a person is. "The experience of will", he writes, "is the way our minds portray their operations to us, not their actual operation" (Wegner, 2005). 'We'

⁴It might be good to mention that this intuitive type of causal inference can, of course, be mistaken: a person might conclude, for instance, that smoking is a cause of good health, because his grandfather died at the respectable age of 93 and has smoked two packs of cigarettes per day ever since he was 17 years old. This causal inference would simply be wrong, but this does nothing to lessen the fact that, in most cases, the inferences *are* correct.

⁵The same goes for these kinds of inferences, although we might expect them to be, in general, more accurate.

as persons thus are the bundle of perceptions designated by the term “us”, distinct from “our minds”. This actually goes beyond a Humean definition of personhood and towards a Cartesian view, which brings with it the problem that it is entirely unclear who the ‘we’ is to whom the mind is portraying its operation.

In a commentary on Wegner, Daniel Dennett makes a similar point about a different remark in Wegner’s text: “we can’t possibly know (let alone keep track of) the tremendous number of mechanical influences on our behavior because we inhabit an extraordinarily complicated machine” (Wegner, 2002). Dennett writes: “who or what is this “we” that inhabits the brain? A Cartesian ghost in the machine? Surely not, in spite of first appearances” (Dennett, 2005).

In order to understand human behavior and the human mind, as well as conscious will, Cartesian-sounding remarks such as Wegner’s should either be justified in some way or avoided altogether, and I think Wegner’s failure to do either of them ultimately leads him to conclude conscious will is an illusion.

The embodied and embedded perspective on cognition considers brain, body, as well as environment vital in the study of organisms and their behavior. It has proven very successful in explaining adaptive behavior (Chiel & Beer, 1997) and I believe one of the lessons that can be learned from its success is that a definition of a person in terms of his or her experiential aspects is a too narrow one; unconscious factors that might not even reside in the brain can prove just as great a contribution to that person’s ‘essence’, i.e., to what makes that person that person.

Dennett (1991) has argued that beliefs, intentions, and other ‘mental’ items should be considered as patterns that can function, as *patterns*, as predictive elements that lawfully covary with some behavioral outcome. This view has the great merit that it takes Hume’s analysis of causality seriously: actions can be considered as caused by patterns, mental or otherwise, if these patterns make some testable difference in the world (i.e., the resulting actions). Dennett rightly does not reserve the status of cause-and-effect pair for items that can directly be mapped onto natural laws discovered by physics (or chemistry).

Some might take terms like ‘belief’ and ‘intention’ to have a certain ‘experiential’ ring to it. This is certainly not my intention: you can believe or intend something while being unaware of it; the beliefs and intentions thus held can still function as predictive patterns and, as such, as causal powers. Of course, such unconscious beliefs and intentions can become, through subsequent self-interpretation (for instance, when asked “what do you think of this?”), conscious ones. This kind of self-interpretation is also crucial for the experience of conscious will and will be treated in the next section.

Venturing further into the exact nature of beliefs, intentions, and other mental items, is beyond the scope of this paper. It suffices to conclude this section by stating that a person can be considered, in a sense that is just as real as causation itself, the cause of some action if, somewhere within that person, a pattern resides that has a predictive, and thus, causal, power over that action.

Experiencing Predictive Patterns as Causal Powers

As already discussed, people are quite good at making causal inferences because being good at this is so hugely convenient. I believe people can be expected to be just as good at making inferences to the conclusion that some predictive pattern within themselves is the cause of their action. The ability to make such inferences would automatically come with a power to predict what you are going to do: making a causal inference is nothing other than establishing a predictive link. Being able to know and talk about what you did and are going to do is probably an enormously advantageous feat in social interaction, and I believe it is this advantage that led to the evolution of the trait.

The experience of conscious will is probably, I would wholly agree with Wegner, due to a process of more or less automatic self-interpretation that follows the principles of priority, consistency, and exclusivity. Causal processes usually do follow these principles⁶, so if you want to determine the cause of something, it would make quite a lot of sense to use those principles in your reasoning.

That the experience of will is a result of a causal inference should mean, in no way, that will is an illusion. The experience of will is simply the perception, or *feeling*, if you will, of the predictive patterns residing within a person and having had, in the (phylogenetic or ontogenetic) past, a causal force over those person’s actions. Of course, the inference can turn out to be wrong, and, in that case, the feeling of will is going to differ from the actual causal process. If, at any given time, there is such a discrepancy, one might be said to have, at that time, an ‘illusory’ experience of will, but when the discrepancy is not there and feeling and actual causal process correspond, the experience should be called veridical.

That a process of perception can be wrong should be no surprise, for it is in no ways limited to the perception of will. Many examples of illusions in visual perception exist. Perhaps the most famous is the Müller-Lyer illusion (see figure 1). Two horizontal line segments are presented to a subject, one of them accompanied by arrowheads pointing towards the line, one of them by arrowheads pointing away from the line. Subjects usually perceive the line segment with arrowheads pointing towards the line as longer, even though the line segments actually are of equal length (Müller-Lyer, 1889). A discrepancy thus exists between the actual state of affairs in the world (i.e., the length of the line segments) and the subjects’ perception of it. Does this possibility of discrepancy mean that visual perception of length is illusory? Should we conclude that ‘conscious length’, as we might call the result of this perception, is an illusion? No, the illusion occurs only in some cases; usually our perception, or experience, of length is the perception of *the actual* length of line segments.

The analogy with visual perception can be taken further. Wegner claims that conscious will is a useful illusion, because the real causes of our behavior are far too complex to keep track of and our minds ‘simplify things a bit’ for us. I would claim, however, that, for instance, visual perception is just as complex as the factors that combine to produce behavior: rays of light reflect off of some object, say, a tree, subsequently

⁶Or do we call them causal precisely because they follow the principles? It does not matter, for even if that were the case, processes we call causal still would obey the principles that ‘define’ them.

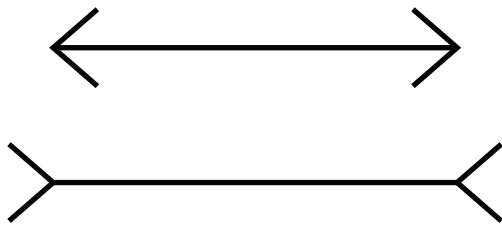


Figure 1: The standard Müller-Lyer illusion.

strike our retina, resulting in frequency-encoded patterns of receptor cell activity that progress, through various brain nuclei, to diverse areas of the brain, both to areas traditionally considered ‘visual areas’, as well as to areas primarily responsible for a host of things seemingly unrelated to visual perception. ‘We’ certainly cannot keep track of this enormous amount of mechanistic interaction and calculation going on. Should we now conclude that our minds “portray to us” that we see a tree and that, in reality, the tree is an illusion? Certainly not: the tree is as real as it can be. Indeed, we perceive it as a tree because that is the way we are ‘wired up’, so to speak, but that is how it should be: we would not be wired up in this way if there were no trees. We experience conscious will because we are wired up to experience it as such; if it were not real we would not be wired up the way we are.

Wegner rightly claims we do not have “direct perception” of the force of will (Wegner, 2002), but I would point out to him that we do not have “direct perception” of anything; we only have ‘mediated perception’ that can be of actual things nonetheless. When one claims that the lack of direct perception of the force of will is reason enough to conclude that will is an illusion, one should, by the same argument, conclude that everything we perceive is, in fact, an illusion created by our minds for the sake of simplifying things a bit ‘for us’. I doubt Wegner is willing to bite that bullet.

Conclusion

Hume’s analysis of causality shows us that any classification of a series of events as causal is always an inference. Because we humans have evolved and grown up in a world that is so highly-causally-seeming, it is not surprising to see we are very good at making such inferences.

Humans often interpret their actions as being caused by themselves. This is also a causal inference: we establish a predictive link between patterns within ourselves and our actions and, as a result, experience conscious will. The experience of conscious will is thus an experience of a causal relation. The ability to establish this experience greatly facilitates social interaction and its acquisition would probably have conferred a significant evolutionary advantage.

Like any causal inference and, indeed, like any perception, the perception of ourselves as having caused some action can be mistaken. However, this is, as I have shown, no reason to

conclude the experience itself is illusory.⁷

References

- Baillargeon, R., Kotovsky, L., & Needham, A. (1995). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack, & A. James-Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 79–116). Oxford, UK: Oxford University Press.
- Baker, L. R. (2007). *The metaphysics of everyday life: An essay in practical realism*. Cambridge University Press.
- Chiel, H., & Beer, R. (1997). The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, 20(12), 553–557.
- Dennett, D. (1984). *Elbow room: The varieties of free will worth wanting*. Cambridge, MA: Bradford Books/MIT Press.
- Dennett, D. (1991). Real patterns. *Journal of Philosophy*, 88(1), 27–51.
- Dennett, D. (2005). Calling in the Cartesian loans. *Behavioral and Brain Sciences*, 27(5), 661.
- Faraday, M. (1853). Experimental investigation of table turning. *Athenaeum*, July, 801–803.
- Hume, D. (1978). *A treatise of human nature*. Oxford, UK: Oxford University Press. (Original work published 1739)
- Hume, D. (2005). *Enquiries concerning human understanding and concerning the principles of morals*. Oxford, UK: Oxford University Press. (Original work published 1777)
- Müller-Lyer, F. C. (1889). Optische Urteilstauschungen. *Dubois-Reymonds Archiv für Anatomie und Physiologie*, 263–270.
- Scepkowski, L., & Cronin-Golomb, A. (2003). The alien hand: Cases, categorizations, and anatomical correlates. *Behavioral and Cognitive Neuroscience Reviews*, 2(4), 261–277.
- Wegner, D. (2002). *The illusion of conscious will*. Cambridge, MA: Bradford Books/MIT Press.
- Wegner, D. (2003). The mind’s best trick: how we experience conscious will. *Trends in Cognitive Sciences*, 7(2), 65–69.
- Wegner, D. (2005). Précis of the illusion of conscious will. *Behavioral and Brain Sciences*, 27(5), 649–659.

⁷One subject that I have left completely untouched in this paper is that of determinism versus indeterminism. I believe the points I have tried to say something about are more important in understanding the human mind and the nature of conscious will, so for the present purposes I will only say that I completely agree with both Dennett (1984) and Wegner in concluding that the kind of free will displayed by Wegner’s example of a radical ‘Free Willer’ that makes decisions “unresponsive to any past influence” (Wegner, 2005, emphasis in original) is certainly not what you would want for free will.