This work originally appeared in:

Dennett, Daniel C. Review of Making Ourselves at Home in Our Machines: The Illusion of Conscious Will, by Daniel Wegner. The Journal of Mathematical Psychology 47, no. 1 (2003): 101-104.

It is available online at:
http://www.sciencedirect.com/science/journal/00222496

ACADEMIC PRESS Journal of Mathematical Psychology 47(1) (2003) 10
Available online at www.sciencedirect.com

Journal of Mathematical Psychology

Book review

Making Ourselves at Home in Our Machines: The Illusion of Conscious Will, Daniel Wegner, MIT Press. Cambridge, MA. 2002, pp. 405, price $34.95, ISBN 0-262-23222-7

The onward march of science and technology makes many thoughtful people uneasy, for good reason. As mysteries are turned into puzzles and then solved, there is less and less room to hide from the dark suspicions that have haunted us since Aristotle's day: are we human beings kidding ourselves when we think we act freely, and for reasons? Are we just "meat machines" whose so-called acts have no more morality attached to them than the destructive "fury" of a tornado, the "gift" of a tree bearing fruit'? The title of Daniel Wegner's book is ominous, and the case he makes for his central claim that conscious will is in one important sense, illusory is unsettling, but his conclusion is not as dire as it first may appear. Conscious will is not at all what we may have thought it was, what tradition supposes it to be, but what it is-or what we have instead of conscious will, if you prefer to let tradition anchor the definition of the terms-is enough to ground our most important ethical convictions, to secure our responsibility for at least many of the things we do. "Illusory or not, conscious will is the person's guide to his or her own moral responsibility for action" (p. 341).

We think we know "from the inside" what we are doing and why, but we also know that there are many things going on in us that we don't have such privileged access to, so how do we know-how could we know that we are actually doing the deciding? Notice how the introduction of the issue of privileged access automatically puts us onto the slippery slope to the Cartesian Theater: the mythical place in the brain "where it all comes together for consciousness" (Dennett, 1991). There are things going on in me that I don't know about, and then there are things I know about "directly"-they are somehow delivered to me wherever I am. Instead of fighting this tempting but treacherous image. Wegner permits himself the full Cartesian picture when it suits his purposes: "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" (p. 27). These machines "we inhabit" simplify things for our benefit: "The experience of will, then, is the way our minds portray their operations to us, not their actual operation" (p. 96). In other words, we get a useful but distorted glimpse of what is going on in our brains:

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, unconscious and inscrutable mechanisms create both conscious thought about action and the action, and also produce the sense of will we experience by perceiving the thought as cause of the action. So, while 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. (p. 98)

Who or what is this "we" that inhabits the brain? It is a commentator and interpreter with limited access to the actual machinery, more along the lines of a press secretary than a president or boss.

In the 18th century. David Hume argued that we never perceive causation directly. What we perceive is succession, first the apparent cause and then the apparent effect, and it is the constant conjunction of similar cause-effect pairs that drives into our minds the idea that there is a necessary-not merely coincidental or contingent--connection between events of the two types. This idea of necessary connection is in some regards illusory: we think we can actually see or observe A causing B, but we never do. Our minds supply the sense of oomph, not the world. Hume's analysis of causation is one of the few success stories in philosophy. In most regards it has stood the test of time remarkably well, and been retroactively supported, you might say, by a host of mundane phenomena. Movies and television wouldn't "work." for instance, if we could-or had to-see causation: the absence of real causation between the image of Bugs Bunny's fist and the image of Elmer Fudd's chin would dispel the illusion that Bugs' hitting Elmer was what caused Elmer to fall over backwards.

Wegner starts from, and expands on, Hume's insights on causation, extending the fundamental message to our knowledge of mental causation the apparent causation of our own deeds by our own decisions or acts of will. We think we know "directly" by some sort of introspection when we act on purpose or intentionally, and we may even suppose that this intimate knowledge we have of our willed actions is somehow immune to

error or tampering. Wegner shows, in many fascinating and delightful ways, that this is simply mistaken. Our access to our "conscious wills" causing our "intentional actions" is fallible. We can be readily fooled because the normal self-knowledge we have is Humean knowledge, just like our knowledge of what causes windows to break when hit by baseballs. William F. Buckley tells the tale in one of his books about friends who lived in an apartment in Paris with a picture window providing a grand view of the Eiffel Tower, which is (or was then) lit up every evening precisely at 6 pm. This couple liked to amuse themselves with a little prank when people came over for drinks and dinner. They kept several precisely set clocks in view, and as 6 o'clock approached, one of them would say "Honey, why don't you turn on the lights," and the other would say, "OK, dear," and walk casually over to the dummy light switch beside the picture window. Five, four, three, two, one, click! To the amazement of the guests, it appeared that these Americans in Paris controlled the floodlights on the Eiffel Tower. Wegner has developed a variety of similar pranks to play on his experimental subjects, testing and refining hypotheses about how their self-knowledge can be manipulated by changing the circumstances in which they act.

One of the phenomena that Wegner exposes for a better view is ideomotor automaticity, the familiar-but always unsettling-phenomenon in which thinking about something can bring about a bodily action related to that thing without the action being an intentional action. For instance, you might betray a secret sexual thought with a tell-tale hand motion that you didn't intend, and in fact would be embarrassed to discover. In such a case, you are not conscious of the causal relation between the thought and the act, but there it is, as good as the causal relation between the aroma of good food and salivation. The main feature of ideomotor actions is people's obliviousness to them-their underprivileged access, you might say. It is as if our usually transparent minds had curtains or barriers installed, behind which these causal chains could get tugged without our introspecting them, producing effects without our compliance. "This ghost army of unconscious actions provides a serious challenge to the notion of an ideal human agent. The greatest contradictions to our ideal of conscious agency occur when we find ourselves behaving with no conscious thought of what we are doing" (p. 157).

For Descartes, the mind was perfectly transparent to itself, with nothing happening out of view, and it has taken more than a century of psychological theorizing and experimentation to erode this ideal of perfect introspectability, which we can now see gets the situation almost backwards. Consciousness of the springs of action is the exception, not the rule, and it requires some rather remarkable circumstances to have evolved at all. Ideomotor actions are the fossils, in effect, of an earlier age, when our ancestors were not as clued in as we are about what they were doing. As Wegner says, "Rather than needing a special theory to explain ideomotor action, we may only need to explain why ideomotor actions and automatism have eluded the mechanism that produces the experience of will" (p. 150).

This mechanism arose as part of the package that evolved in our species along with language. "A voluntary action is something a person can do when asked," Wegner notes (p. 32), and this quite sharply distinguishes human action from animal action. When psychologists and neuroscientists devise a new experimental setup or paradigm in which to test non-human subjects such as rats or cats or monkeys or dolphins, they often have to devote dozens or even hundreds of hours to training each subject on the new tasks. A monkey, for instance, can be trained to look to the left if it sees a grating moving up and look to the right if it sees a grating moving down. All this training takes time and patience, on the part of both trainer and subject. Human subjects in such experiments, however, can usually just be told what is desired of them. After a brief question-and-answer session and a few minutes of practice, we human subjects will typically be as competent in the new environment as any agent ever could be. Of course, we do have to understand the representations presented to us in these briefings, and what is asked of us has to be composed of action-parts that fall within the range of things we can do. That is what Wegner means when he identifies voluntary actions as things we can do when asked. If asked to lower your blood pressure or adjust your heartbeat or wiggle your ears, you will not be so ready to comply, though with training not unlike that given to laboratory animals, you may eventually be able to add such feats to your repertoire of voluntary actions.

When language came into existence, it brought into existence the kind of mind that can transform itself on a moment's notice into a somewhat different virtual machine, taking on new projects, following new rules, adopting new policies. We are transformers. That's what a mind is, as contrasted with a mere brain: the control system of a chameleonic transformer. A virtual machine for making more virtual machines. Non-human animals can engage in voluntary action of sorts. The bird that flies wherever it wants is voluntarily wheeling this way and that, voluntarily moving its wings, and it does this without benefit of language. The distinction embodied in anatomy between what it can do voluntarily (by moving its striated muscles) and what happens autonomically, moved by smooth muscle and controlled by the autonomic nervous system, is not at issue. We have added a layer on top of the bird's (and the ape's and the dolphin's) capacity to decide what to do next. It is not an anatomical layer in the brain, but a functional layer,

a virtual layer composed somehow in the micro-details of the brain's anatomy: We can ask each other to do things, and we can ask ourselves to do things. And at least sometimes we readily comply with these requests. Yes, your dog can be "asked" to do a variety of voluntary things, but it can't ask why you make these requests. A male baboon can "ask" a nearby female for some grooming, but neither of them can discuss the likely outcome of compliance with this request, which might have serious consequences for both of them, especially if the male is not the alpha male of the troop. We human beings not only can do things when requested to do them: we can answer inquiries about what we are doing and why. We can engage in the practice of asking, and giving, reasons.

It is this kind of asking, which we can also direct to ourselves, that creates the special category of voluntary actions that sets us apart. Other, simpler intentional systems act in ways that are crisply predictable on the basis of beliefs and desires we attribute to them on the basis of our surveys of their needs and their history, their perceptual and behavioral talents, but with us, it is different. We need to have something to say when asked what the heck we think we're doing. And when we answer, our authority is problematic. The evolutionary biologist William Hamilton. reflecting on his own uneasiness with his recognition of this fact, put the issue particularly well:

In life, what was it I really wanted? My own conscious and seemingly indivisible self was turning out far from what I had imagined and I need not be so ashamed of my self-pity! I was an ambassador ordered abroad by some fragile coalition, a bearer of conflicting orders, from the uneasy masters of a divided empire. As I write these words, even so as to be able to write them, I am pretended to a unity that, deep inside myself, I now know does not exist. (Hamilton, 1996, p. 134)

Once language evolved, people could do things with words that they could never do before, and the beauty of the whole development was that it tended to make those features of their complicated neighbors that they were most interested in adjusting readily accessible to adjustment from outside-even by somebody who knew nothing about the internal control system, the brain. These ancestors of ours discovered whole generative classes of behaviors for adjusting the behavior of others, and for monitoring and modulating (and if need be resisting) the reciprocal adjustment of their own behavioral controls by those others.

The centerpiece metaphor of this co-evolved human user-illusion is the Self, which appears to reside in a place in the brain, the Cartesian Theater, providing a limited, metaphorical outlook on what's going on in our brains. It provides this outlook to others, and to ourselves. In fact, we wouldn't exist, as Selves "inhabiting complicated machinery" as Wegner so vividly puts it, if it weren't for the evolution of social interactions requiring each human animal to create within itself a subsystem designed for interacting with others. Once created, it could also interact with itself at different times. Until we human beings came along, no agent on the planet enjoyed the curious non-obliviousness we have to the causal links that emerged as salient once we human beings began to talk about what we were up to. As Wegner puts it, "People become what they think they are, or what they find that others think they are, in a process of negotiation that snowballs constantly" (p. 314).

Wegner is right, then, to identify the Self that emerges in his and others' experiments as a sort of public-relations agent, a spokesperson instead of a boss, but these are extreme cases set up to isolate factors that are normally integrated, and we need not identify ourselves so closely with such a temporarily isolated self. Wegner draws our attention to the times-not infrequent among those of us who are "absent-minded" when we find ourselves with a perfectly conscious thought that just baffles us: it is, as he wonderfully puts it, conscious but not accessible (p. 163). (Now why am I standing in the kitchen in front of the cupboard? I know I'm in the place I meant to be, but what did I come in here to get?) At such a moment, I have lost track of the context, and hence the raison d'etre, of this very thought, this conscious experience, and so its meaning (and that's what is most important) is temporarily no more accessible to me-the larger me that does the policy-making than it would be to any third party, any "outside" observer who came upon it. In fact, some onlooker might well be able to remind me of what it was I was up to. My capacity to be reminded (re-minded) is crucial, since it is only this that could convince me that this onlooker was right, that this was something I was doing. If the thought or project is anyone's, it is mine it belongs to the me who set it in motion and provided the context in which this thought makes sense: it is just that the part of me that is baffled is temporarily unable to gain access to the other part of me that is the author of this thought.

I might say. in apology, that I was not myself when I made that mistake, or forgot what I was about, but this is not the severe disruption of self-control that is observed in schizophrenia, in which the patients own thoughts are interpreted as alien voices. This is just the fleeting loss of contact that can disrupt a perfectly good plan. A lot of what you are, a lot of what you are doing and know about, springs from structures down there in the engine room, causing the action to happen. If a thought of yours is only conscious, but not also accessible to that machinery (to some of it, to the machinery that needs it), then you can't do anything

with it, and are left just silently mouthing the damn phrase to yourself, your isolated self, over and over. Isolated consciousness can indeed do nothing much on its own. Nor can it be responsible.

As Wegner notes, "If people will often forget tasks for the simple reason that the tasks have been completed, this signals a loss of contact [emphasis added] with their initial intentions once actions are over-and thus a susceptibility to revised intentions" (p. 167). A loss of contact between what and what? Between a Cartesian Self that "does nothing" and a brain that makes all the decisions? No. A loss of contact between the you that was in charge then and the you that is in charge now. A person has to be able to keep in contact with past and anticipated intentions, and one of the main roles of the brain's user-illusion of itself, which I call the self as a center of narrative gravity, is to provide me with a means of interfacing with myself at other times. As Wegner puts it, "Conscious will is particularly useful, then, as a guide to ourselves"(p. 328). The perspectival trick we need in order to escape the clutches of the Cartesian Theater is coming to see that I, the larger, temporally and spatially extended self, can control, to some degree, what goes on inside of the simplification barrier, where the decision-making happens, and that is why, as Wegner says, "Illusory or not, conscious will is the person's guide to his or her own moral responsibility for action" (p. 341).

This book is what books are often called and seldom are: required reading for anybody who wants to think about free will, a classic in the making. Aren't we lucky, then, that it is also a joy to read, with many amusing touches and vivid formulations.[1]

References

Dennett, D. C. (1991). Consciousness explained. Boston: Little, Brown.
Dennett, D. C. (2003). Freedom evolves. New York: Viking Penguin.
Hamilton, W. (1996). Narrow roads of gene land' Vol 1. Evolution of social behaviour. Oxford: Freeman.

Daniel C. Dennett
Philosophy Department, Tufts University,
Medford, MA 02155, USA

Daniel C. Dennett is University Professor and Austin B. Fletcher Professor of Philosophy, and Director of the Center for Cognitive Studies at Tufts University. He received his B.A. in philosophy from Harvard in 1963 and the D.Phil. in philosophy from Oxford in 1965. He taught at U.C. Irvine from 1965 to 1971, when he moved to Tufts. He is the author of nine books, among them Consciousness Explained (1991), Darwin's Dangerous Idea (1995), and Freedom Evolves (2003), as well as many articles on mind, evolution, consciousness and artificial intelligence. Currently he is working on conceptual problems confronting empirical theories of human consciousness, and the implications of these problems for social issues.

[1] Portions of this review are drawn from Dennett, 2003, Chapter 8, with revisions.