The mind's best trick: how we experience conscious will

Daniel M. Wegner

Department of Psychology, Harvard University, 1470 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA

We often consciously will our own actions. This experience is so profound that it tempts us to believe that our actions are caused by consciousness. It could also be a trick, however – the mind's way of estimating its own apparent authorship by drawing causal inferences about relationships between thoughts and actions. Cognitive, social, and neuropsychological studies of apparent mental causation suggest that experiences of conscious will frequently depart from actual causal processes and so might not reflect direct perceptions of conscious thought causing action.

Does consciousness cause action? Many people think that even asking this question is absurd. How could consciousness *not* cause what we do? Every few moments of every day, we think about doing something and then do it. We think of moving a finger and then do it, we think of going to the store for milk and do it, we think of looking away from this page – and then do it. It certainly doesn't take a rocket scientist to draw the obvious conclusion from a lifelong accumulation of such examples: consciousness is an active force, an engine of will.

The mind has been known to play tricks, though. Could this be one? What if our minds keep showing us the same set of appearances, leading to an impression of conscious will again and again, but never revealing to us how our actions are actually caused? One way this could happen is if both the thought about action and the action itself are caused by unperceived forces of mind: you think of doing X and then do X – not because conscious thinking causes doing, but because other mental processes (that are not consciously perceived) cause both the thinking and the doing. Based on your conscious perceptions of your thoughts and actions, it would be impossible to tell in any given case whether your thought was causing your action, or something else was causing both of them. Could it be that the deep intuition we all have about the power of our conscious will is the result of this 'sleight of mind'? Perhaps we experience conscious will when we infer that our thought causes our action, although we can't really know that this is the causal path (see Fig. 1).

Anomalies of will

If conscious will were an illusory add-on to action, we could begin to explain all the odd cases when action and conscious will do not properly coincide.

Neuropsychology

We might understand, for example, Penfield's classic finding on movements induced through electrical stimulation of the motor cortex [1]. Conscious patients were prompted by stimulation of the exposed brain to produce movements that were not simple reflexes and instead appeared to be complex, multi-staged, and voluntary. Yet, their common report of the experience was that they did not 'do' the action, and instead felt that Penfield had 'pulled it out' of them. This observation only makes sense if the experience of will is an addition to voluntary action, not a cause of it.

The possibility that conscious will is an illusion might also explain the finding that unperceived causes of action can fail to influence the experience of will [2]. People in one study, for instance, were asked to choose to move one or the other index finger whenever they heard a click [3]. Transcranial magnetic stimulation (TMS) was applied alternately to the left or right motor cortex to influence the movement, and this influence over which finger was moved was strong at short response times. Respondents reported consciously willing the movements during the TMS influence, although showing a lack of insight into the alternative causal mechanism producing their actions. Similar inferences can be drawn from Gazzaniga's observations of split-brain patients who are induced to perform an action through communication to the right hemisphere when the major verbal centers of the left hemisphere are unaware of the action's cause [4]. Such patients confabulate 'left brain interpretations' of their intentions, apparently to satisfy the general assumption that their actions are consciously willed.

The celebrated experiments of Benjamin Libet provide further evidence that conscious will can be experienced that does not correspond to causation [5]. In spontaneous, intentional finger movement, Libet found that a scalp-recorded brain readiness potential (RP) preceded the movement (measured electromyographically) by a minimum of ~550 ms. This finding indicates only that some sort of brain activity reliably precedes the onset of voluntary action. However, participants were also asked to recall the position of a clock at their initial awareness of intending to move their finger, and this awareness *followed* the RP by

Corresponding author: Daniel M. Wegner (wegner@wjh.harvard.edu).

http://tics.trends.com 1364-6613/03/$ - see front matter doi:10.1016/S1364-6613(03)00002-0

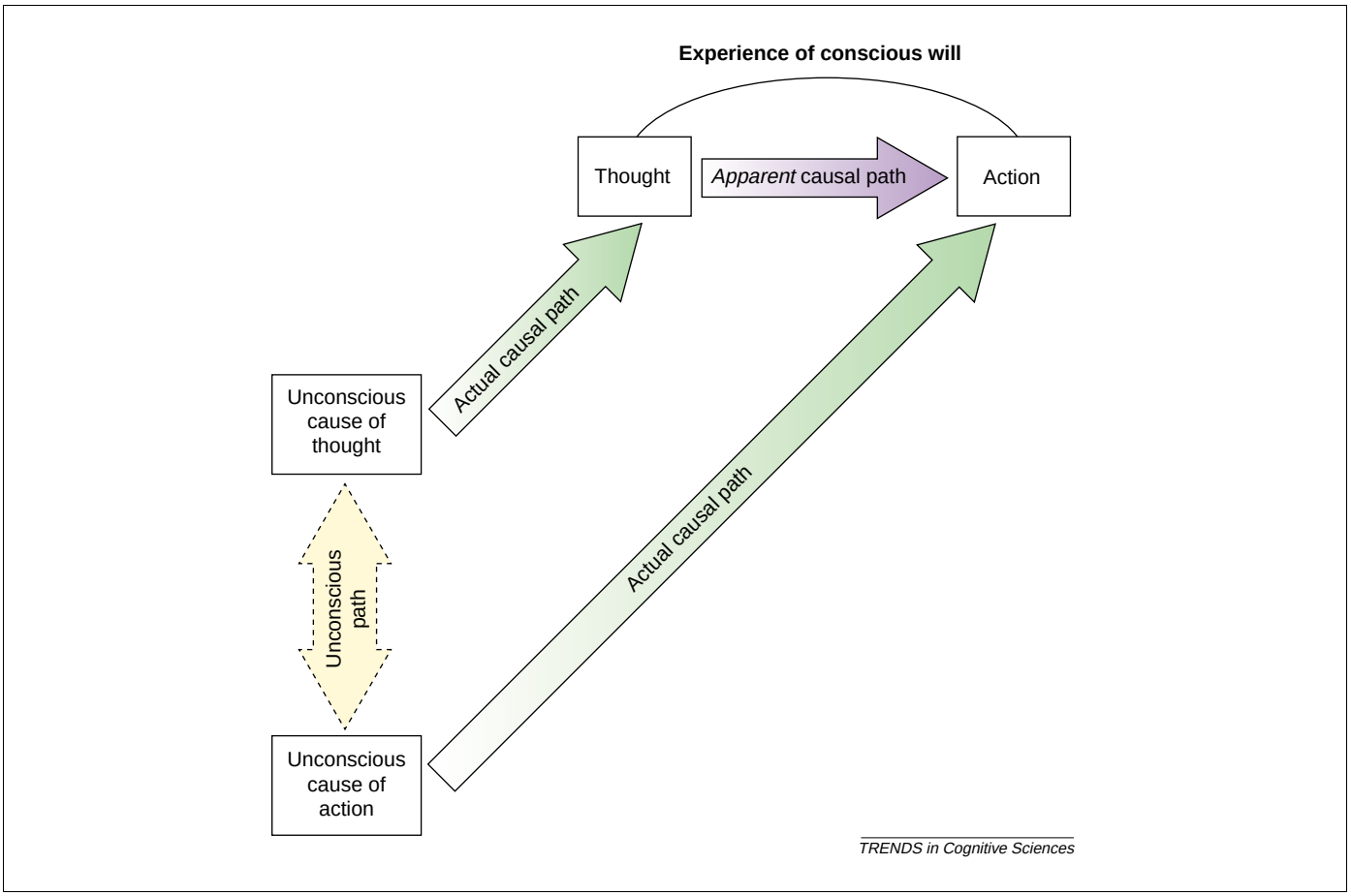

Fig. 1. The experience of conscious will arises when the person infers an apparent causal path from thought to action (purple arrow). The actual causal paths (green) are not present in the person's consciousness. The thought is caused by unconscious mental events, and the action is caused by unconscious mental events, and these unconscious mental events might also be linked to each other directly or through yet other mental or brain processes. Conscious will is experienced as a result of what is apparent, not what is real. Modified with permission from Ref. [22].

some 350–400 ms. So, although the conscious intention preceded the finger movement, it occurred well after whatever brain events were signaled by the RP. This finding suggests that the experience of consciously willing an action begins after brain events that set the action into motion [6,7]. The brain creates both the thought and the action, leaving the person to infer that the thought is causing the action.

Clinical evidence

Anomalies pointing to a system that fabricates an experience of will can also be found in clinical cases. Patients with brain damage resulting in 'alien hand syndrome', for example, report that one of their hands functions with a mind of its own, often performing elaborate and seemingly voluntary actions without the patient's experience of willful control. One patient described the experience as a feeling that 'someone from the moon' was controlling her hand [8]. Schizophrenia accompanied by auditory hallucinations also produces anomalistic will – in this case, an experience of 'hearing voices' that occurs when patients attribute their own thoughts and inner voice to others [9–14]. Thoughts that come to mind without prior anticipation are not experienced as willed, and their insistent recurrence can lead patients to ascribe them to outside agents.

Automatisms

Will is also experienced independently of action in a menagerie of cases known as automatisms [15–19]. Practices such as automatic writing, table turning, Ouija-board spelling, dowsing, pendulum divining, channeling, and the like were the major psychological basis of the Spiritualist fad of the late 19th century, as these various contrivances gave rise to experiences of unwilled action that were then attributed to spirits or other supernatural agents. In the case of table turning, for instance, a group of people gathered around a light table and waited for it to move (Fig. 2). Often it would – after a significant wait – sometimes even circling the room or rocking from side to side. Yet the participants often reported no experience of willing the action and instead expressed amazement at the table's animation. Although spirit agency was the popular explanation, investigations by scientists such as Michael Faraday (using force measurement devices between hands and tables) revealed that the source of the table movement was indeed the participants [20,21]. The experience of will in such cases was entirely misleading about the causal basis of the action.

Apparent mental causation

If the experience of conscious will is not a direct report

Fig. 2. Parisians in 1853 test the automatic (i.e. unwilled) turning of a piano stool. This kind of practice was part of the Spiritualist fad in the 19th century. Reproduced from Ref. [21].

of the processes whereby action is produced, what is it? The likely sources of the experience of conscious will are the topic of the 'theory of apparent mental causation' [19,22].

Principles

This theory suggests that conscious will is experienced when we draw the inference that our thought has caused our action – whether or not this inference is correct. The inference occurs in accordance with principles that follow from research on cause perception and attribution [23–27] – principles of priority, consistency, and exclusivity. When a thought appears in consciousness just before an action (priority), is consistent with the action (consistency), and is not accompanied by conspicuous alternative causes of the action (exclusivity), we experience conscious will and ascribe authorship to ourselves for the action. In essence, we experience ourselves as agents who consciously cause our actions when our minds provide us with timely previews of actions that turn out to be accurate when we observe the actions that ensue. Elements of this theory can be traced to David Hume, and can be understood as an application of his general analysis of the perception of causality [28].

In commonplace actions, we often have thoughts of action that are consistent, prior, and exclusive. We think of turning on a light before doing so, for example, and nothing else seems to be causing the light to go on, so when it happens we conclude that we did it. If we were not thinking of turning on the light and found ourselves flipping the switch, the lack of consistency between our thought and action might undermine the feeling of conscious will for the action. If we thought of turning on the light only *after* the action, the lack of appropriate priority could also yield little sense of will for the action. And if we saw someone's hand on a parallel switch at the other door to the room, we might be inclined to feel less conscious will for the light's turning on even if we had thought of doing it just in advance of our own movement.

Illusory will

Imagine for a moment that a consistent thought occurs before an action, and that the thought is not accompanied by any other potential causes of action – but that it does *not* cause the action. One might still perceive it as causal nonetheless. On thinking of the light switching on just before it actually comes on by another cause, for example, one might have the distinct but erroneous impression that one had caused it. The principles of causal inference do not describe actual causal relations, then, because the perceptions of causality that derive from the principles can depart from reality [29]. When inferences diverge from actual causal sequences, the experience of conscious will goes awry, leaving the person to experience authorship of actions that could not have been theirs, or that may not even have occurred – or on the other hand to experience no authorship even when the action is demonstrably linked to the person [30]. Departures from veridical perceptions of causality can be traced to each of the principles.

The consistency principle, for example, suggests the general proposition that people will feel more will for success than failure. After all, people more often envision success of a task than failure, so when success occurs, the consistency between the prior thought and the observed action produces an experience of will. Such effects have been observed in studies of the perceived contingency between actions and outcomes. People perceive that they controlled a chance event when they receive a large number of initial successes in predicting that event [31]. The perception that one is causing a successful outcome is enhanced merely by the increased frequency of that outcome [32]. This also makes sense of the fact that depressed individuals – who think less often of success – are not as likely as others to over-perceive control of successful outcomes [33].

The priority principle carries with it further implications for the experience of will. It suggests, for instance, that people will think they have caused actions when a thought relevant to the action is primed just before the action – whether they actually performed the action or not. People in one experiment were presented with thoughts (e.g. a tape-recorded mention of the word *swan*) relevant to their action (moving an onscreen cursor to select a picture of a swan) [22]. The movement that participants performed was not in fact their own, as they shared the computer mouse with an experimental confederate who gently forced the action without the participants' knowledge. (On other trials, the effect of the thought on the participant's own action was found to be nil when the

action was not forced.) Nevertheless, when the relevant thought was provided either 1 s or 5 s before the action, participants reported feeling that they acted intentionally in making the movement. The influence of priority was shown to be present because, on trials when thoughts of the swan were prompted 30 s before the forced action or 1 s afterwards, no inflated experience of will was found. Even when the thought of the action is wholly external – as in this case, presented over headphones – its timely appearance before the action led to an enhanced experience of apparent mental causation.

The exclusivity principle governs cases when perceptions of forces outside the self undermine the experience of will. The presence of other actors who could contribute to the action, for instance, creates circumstances in which people can fail to sense willful control of their actions. In the case of hypnosis, the experience of performing suggested actions gives many susceptible individuals strong impressions of involuntariness despite their obvious involvement in acting [34]. Stanley Milgram explained his famous finding, that people will obey a command to shock another person, in terms of such a mechanism, suggesting that an 'agentic shift' and an accompanying reduction in conscious will occur when actions are done at the behest of another [35].

At the extreme, perceptions of other agents can obliterate the experience of authorship of one's own action. In a series of studies, people who were asked to sense the detectable muscle movements of another person and then type answers to questions for that person were found to produce correct answers and attribute a substantial contribution to the other, even when the other was an experimental confederate who never heard the questions [19]. This phenomenon appears to underlie facilitated communication, the discredited technique whereby people giving manual support to autistic or other communication-impaired individuals create typed messages that they erroneously attribute to those individuals [36]. The belief in outside agents who influence a person's actions can so muddle the perception of conscious will as to promote bizarre dissociations of perceived authorship in the form of trance channeling, spirit possession, and dissociative identity disorder [19].

Box 1. Questions for future research

- How do the consistency, priority and exclusivity principles interact with each other to influence conscious will?
- Does the experience of conscious will operate by the same processes *before* action as *after* action?
- What brain activities are involved in the experience of conscious will – beyond the activities associated with the production of intention and the production of action?
- Are there ways to enhance self-insight by increasing the correspondence between apparent mental causation and actual causal processes?

Conclusions

Does all this mean that conscious thought does not cause action? It does not mean this at all. The task of determining the causal relations between conscious representations and actions is a matter of inspection through scientific inquiry, and reliable connections between conscious thought and action can potentially be discerned by this process [37]. The point made here is that the mind's own system for computing these relations provides the person with an experience of conscious will that is no more than a rough-and-ready guide to such causation, one that can be misled by any number of circumstances that render invalid inferences (see also Box 1). We should be surprised, after all, if cognitive creatures with our demonstrably fallible self-insight were capable of perceiving the deepest mechanisms of our own minds [38,39]. The experience of conscious will is a marvelous trick of the mind, one that yields useful intuitions about our authorship – but it is not the foundation for an explanatory system that stands outside the paths of deterministic causation.

Acknowledgements

Thanks are due to John Bargh, Celeste Beck and Jonathan Schooler for helpful comments on an earlier draft. Some of the research described here was supported by NIMH Grant 49127.

References

1 Penfield, W. (1975) *The Mystery Of Mind*, Princeton University Press
2 Ammon, K. and Gandevia, S.C. (1990) Transcranial magnetic stimulation can influence the selection of motor programmes. *J. Neurol. Neurosurg. Psychiatr.* 53, 705–707
3 Brasil-Neto, J.P. *et al.* (1992) Focal transcranial magnetic stimulation and response bias in a forced-choice task. *J. Neurol. Neurosurg. Psychiatr.* 55, 964–966
4 Gazzaniga, M.S. (1995) Consciousness and the cerebral hemispheres. In *The Cognitive Neurosciences* (Gazzaniga, M.S., ed.), pp. 1391–1400, MIT Press
5 Libet, B. (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529–566
6 Haggard, P. *et al.* (2002) Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385
7 Haggard, P. and Eimer, M. (1999) On the relation between brain potentials and the awareness of voluntary movements. *Exp. Brain Res.* 126, 128–133
8 Geschwind, D. *et al.* (1995) Alien hand syndrome: interhemispheric motor disconnection due to a lesion in the midbody of the corpus callosum. *Neurology* 45, 802–808
9 Daprati, E. *et al.* (1997) Looking for the agent: an investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition* 65, 71–86
10 Frith, C.D. and Done, D.J. (1989) Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychol. Med.* 19, 359–363
11 Graham, G. and Stephens, G.L. (1994) Mind and mine. In *Philosophical Psychology* (Graham, G. and Stephens, G.L., eds) pp. 91–109, MIT Press
12 Hoffmann, R.E. (1986) Verbal hallucinations and language production processes in schizophrenia. *Behav. Brain Sci.* 9, 503–548
13 Blakemore, S.J. *et al.* (2002) Abnormalities in the awareness of action. *Trends Cogn. Sci.* 6, 237–242
14 Spence, S.A. (1996) Free will in the light of neuropsychiatry. *Philos. Psychiatr. Psychol.* 3, 75–90
15 Ansfield, M. and Wegner, D.M. (1996) The feeling of doing. In *The Psychology of Action: Linking Cognition and Motivation to Behavior* (Gollwitzer, P.M. and Bargh, J.A., eds) pp. 482–506, Guilford Press

16 Bargh, J.A. (1997) The automaticity of everyday life. In *Advances in Social Cognition* (Vol. 10) (Wyer, R.S., ed.), pp. 1–62, Erlbaum
17 Carpenter, W.B. (1875) *Principles of Mental Physiology*, D. Appleton & Company
18 Spitz, H.H. (1997) *Nonconscious Movements: From Mystical Messages to Facilitated Communication*, Erlbaum
19 Wegner, D.M. (2002) *The Illusion of Conscious Will*, MIT Press
20 Faraday, M. (1853) Experimental investigation of table turning. *Athenaeum (July)*, 801–803
21 In *L'Illustration* (1853), p. 325
22 Wegner, D.M. and Wheatley, T.P. (1999) Apparent mental causation: sources of the experience of will. *Am. Psychol.* 54, 480–492
23 Einhorn, H.J. and Hogarth, R.M. (1986) Judging probable cause. *Psychol. Bull.* 99, 3–19
24 Gilbert, D.T. (1995) Attribution and interpersonal perception. In *Advanced Social Psychology* (Tesse, A., ed.), pp. 98–147, McGraw-Hill
25 Kelley, H.H. (1972) Causal schemata and the attribution process. In *Attribution: Perceiving the Causes of Behavior* (Jones, E.E. *et al.*, eds), pp. 151–174, General Learning Press
26 McClure, J. (1998) Discounting causes of behavior: are two reasons better than one? *J. Pers. Soc. Psychol.* 74, 7–20
27 Michotte, A. (1963) *The Perception of Causality*, Basic Books
28 Hume, D. (1888) *A Treatise of Human Nature*, Oxford University Press
29 Kelley, H.H. (1980) Magic tricks: The management of causal attributions. In *Perspectives on Attribution Research and Theory: The Bielefeld Symposium* (Gurlitz, D., ed.), pp. 19–35, Ballinger
30 Nisbett, R.E. and Wilson, T.D. (1977) Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* 84, 231–259
31 Langer, E.J. and Roth, J. (1975) Heads I win, tails it's chance: the illusion of control as a function of the sequence of outcomes in a purely chance task. *J. Pers. Soc. Psychol.* 32, 951–955
32 Jenkins, H.M. and Ward, W.C. (1965) Judgments of contingency between responses and outcomes. *Psychol. Monogr.* 79 No. 1
33 Alloy, L.B. and Abramson, L.Y. (1979) Judgment of contingency in depressed and nondepressed students: sadder but wiser? *J. Exp. Psychol. Gen.* 108, 441–485
34 Lynn, S.J. *et al.* (1990) Hypnotic involuntariness: a social cognitive analysis. *Psychol. Rev.* 97, 169–184
35 Milgram, S. (1974) *Obedience to Authority*, Harper & Row
36 Jacobson, J.W. *et al.* (1995) A history of facilitated communication: science, pseudoscience, and antiscience. *Am. Psychol.* 50, 750–765
37 Jack, A.I. and Shallice, T. (2001) Introspective physicalism as an approach to the science of consciousness. In *The Cognitive Neuroscience of Consciousness* (Dehaene, S., ed.), pp. 161–196, MIT Press
38 Schooler, J.W. (2002) Re-representing consciousness: dissociations between experience and meta-consciousness. *Trends Cogn. Sci.* 6, 339–344
39 Wilson, T.D. (2002) *Strangers to Ourselves: Discovering the Adaptive Unconscious*, Harvard University Press