

Causation and the Awareness of Agency

Holly Andersen

History and Philosophy of Science, University of Pittsburgh

1017 Cathedral of Learning, Pittsburgh PA, 15260

hka1@pitt.edu

Abstract: I criticize the tendency to address the causal role of awareness in agency in terms of the awareness *of* agency, and argue that this distorts the causal import of experimental results in significant ways. I illustrate, using the work of Shaun Gallagher, how the tendency to focus on the awareness of agency obscures the role of extrospective awareness by considering it only in terms of what it contributes to the awareness of agency. Focus on awareness of agency separates awareness from agency itself, and then turns it inwards to introspect distinct agentive processes. If we then assume that the causal influence of awareness is directed at the same object as awareness itself, then the only avenue for conscious causal involvement in action is to somehow interfere with the separate, even neuronal, processes leading to action. I label this the Micromanagement Model of conscious agency, because it forces awareness to micromanage other, nonconscious, processes in order to be causally efficacious. Implicit adherence to the Micromanagement Model prejudices us towards the mistaken conclusion that awareness has limited to no causal role in action.

Introduction

Based on new research in cognitive science in the last decade, a number of philosophers have concluded that we have little to no conscious causal influence on our own actions, and that our strong impressions to the contrary are simply persistent illusions. These strong conclusions are symptoms of a problematic view of conscious action, and it is this problematic view of how awareness might figure in action that is the target of this paper. My focus here is on the model of agency that underlies a fairly widespread tendency to take *awareness of* agency (alternatively, an experience, sense, or phenomenology of agency) as the paradigm for understanding conscious contributions to agency. Regardless of whether or not we endorse the problematic symptoms of this view (such as concluding that conscious awareness has no causal role in action), this model is agency is, I'll show, dubious.

A number of authors utilize the sense or awareness of agency as a method for investigating the extent of the conscious causal contributions to action (Metzinger 2006; Velmans 2004; O'Shaughnessy 2003; Carruthers 2007; Nahmias 2005; Choudhury and Blakemore 2006; Wegner and Wheatley 1999; Haggard 2003). Two manifestations of this trend are, on the one hand, the notion that conscious awareness is involved in action as awareness *of* acting; and, on the other, that introspective reports by agents should reveal awareness of the mechanisms involved in the exercise of agency, if the agent was consciously involved in that exercise (Mossel 2005; Bayne and Levy 2006; Horgan, Tienson, and Graham 2003).¹

This focus has the effect of separating awareness from other, distinct, agentic processes, and then directing awareness at those processes. It effectively reduces the potential role of conscious awareness to that of a passive monitoring mechanism, producing introspective reports on already ongoing processes that are by themselves sufficient to lead to action. In light of such assumptions, the results of experiments such as Libet (1985) shouldn't really surprise us, since it is in the nature of such a monitoring mechanism to produce reports after the fact. Underlying the focus on the awareness of agency and assumption that introspective reports should reveal the mechanisms of action is a flawed model of agency that I will call the Micromanagement Model. On this model, awareness is introspectively directed at distinct agentic processes leading to action, and in order to exert any causal influence in action, awareness must somehow “fiddle with the knobs” of those other agentic processes. The causal role that the Micromanagement Model makes available for awareness is that of intervening in the nonconscious processes that lead to action. Holding awareness to such a skewed and strong standard to count as causally efficacious in action makes it overwhelmingly easy to conclude that our conscious selves play little to no causal role in producing our actions, whether or not actually endorses that view.

¹ There are philosophers (*inter alia*, Chalmers 1996, Newell 1992) who take there to be an important difference between awareness and conscious awareness, whereas I do not distinguish between these (and specifically avoid use of the term ‘consciousness’). For reasons of space, I will not make a philosophical defense here of my usage of these terms. It’s important to note, though, that in the scientific literature on which I rely, awareness and conscious awareness are used interchangeably. As my goal is provide an empirically grounded account of conscious agency, the extant science should guide usage of these terms.

The focus on the sense of agency has three specific shortcomings to be explored in this paper. First and most importantly, it directs the causal influence of awareness inward, towards our own internal processes leading to movement. This internal re-direction separates awareness from agency and constitutes the fundamental error that gives rise to the Micromanagement Model of conscious agency. Second, while the sense of agency may allow us to infer the existence of conscious agency, we have reason to think that the processes leading to movement cause both movement and the sense of agency itself, and as such it could not be a causal factor in generating that movement. If we consider extrospective awareness only in terms of what it contributes to the sense of agency, we have essentially eliminated the substantive avenues by which awareness could causally contribute to action. Finally, the focus on sense of agency ties the causal role of conscious awareness too closely to the ability to provide introspective reports regarding one's own agency, making it a necessary condition for awareness having been causally involved that we be aware of the separate process leading to action.

My claim is neither that there is nothing worthwhile to be found by looking at the sense of agency, nor that all the authors who discuss it also endorse the Micromanagement Model. My argument is that the level to which we have focused on the sense of agency obscures the causal role of nonintrospective awareness in action. The MM need not be a position endorsed as such by any given philosopher, but I am claiming that it is the underlying model that goes along with the sense of agency as a means of investigating conscious action, and that it is what provides apparent plausibility to many of the existing positions regarding conscious agency.

Section 1 illustrates how the focus on awareness of agency distorts the causal role of awareness in action by assimilating nonintrospective aspects of awareness to the sense of agency; I do this by looking a representative case in the work of Shaun Gallagher. Section 2 generalizes the focus on sense of agency and reliance on introspective reports, to demonstrate how such a focus leads to the Micromanagement Model of conscious agency.

Section 1

In this section, I want to show how a focus on sense of agency is so strong that it leads to assimilation of other world-directed, extrospective, aspects of awareness to the sense of awareness. The extrospective aspects become relevant only insofar as they contribute to the awareness of agency, and not as potential causal factors in their own right. This segues in to section 2 by demonstrating how such a focus turns awareness inwards towards one's own, sometimes nonconscious, agentive processes.

Recent work by Shaun Gallagher emphasizes the role of the sense of agency in action by developing a distinction between the sense of ownership of action (that it is oneself which moves or thinks) and the sense of agency (that one is the originator or source of this movement or thinking). For Gallagher, the sense of agency, the feeling that one is the causal source of an action, is closely tied to the immunity principle, that we are immune to making identificational errors, either in propositions or more broadly in experience, that involve attributions to ourselves as a first person pronoun (Gallagher 2000, 4). This connection draws the sense of agency into extremely close contact with introspective reports of the experience of acting. In one regard, that closeness of connection is appropriate: the sense of agency is a kind of feeling one has about oneself, a feeling on which one can provide reports. The problem is then going on to use the sense of agency to stand in for agency itself. The sense of agency is itself compatible with either the presence or absence of conscious causal involvement in the action. The very way in which the distinction between sense of ownership and of agency is framed separates the awareness of agency from the object of that awareness, agency itself. This distinction first separates awareness from agency, which is the basic move that leads to the Micromanagement Model.

Gallagher (2007) states that it is the sense of agency that gives us insight into conscious causal contributions to action. His aim "is to investigate both the phenomenology and science of agency. In its proper sense, I understand agency to depend on the agent's consciousness of agency" (Gallagher 2007, 2). He focuses on the sense of agency as the way to understand conscious contributions to actions, even though he also mentions other aspects of conscious involvement in action beyond the sense of agency:

...clearly our sense of agency for the action will be tied to that intentional aspect, and that aspect is where our attention is directed – *in the world*, in the project or task that we are engaged in. So clearly a form of *intentional feedback*, which is not afferent feedback about our bodily movements, but some perceptual sense that my action is having an effect, must contribute to the sense of agency.

I suggest, then, that the sense of agency, at the first-order level of experience, is complex because it is the product of several contributory elements: efferent signals, sensory (afferent) feedback, and intentional (perceptual) feedback. (Gallagher 2007, 8)

There are two key ideas in this passage. The first is the mention of distinct ways in which conscious awareness could be involved in action; in particular, attention is directed out into the world, not merely at our own bodily movements and sensory feedback. This direction of awareness out into the world is distinct from our awareness of agency, and it is precisely what is neglected by the focus on the sense of agency. That neglect is evident in this passage, because although he brings up these other ways in which awareness is involved in action, Gallagher considers them only in terms of how they relate to the sense of agency. Instead of looking at the causal involvement of world-directed or extrospective awareness itself, he considers world-directed awareness only in terms of what it contributes to the sense of agency. Gallagher ends up ignoring the possibility that extrospective awareness may have a causal contribution to action distinct from its contribution to the sense of agency. Assimilating other aspects of awareness to phenomenology of agency is problematic because the very way in which the experience of agency is separated from agency removes it from the causal processes leading to action. This separation of experience of agency from agency *per se* deprives conscious awareness any real causal role in that agency itself aside from what could be exercised *on* that agency by awareness.

This passage also illustrates why focusing on the awareness of agency distorts the picture we get of the causal contribution of awareness to action. The sense of agency likely arises from, or is strongly influenced by, the same neural processes that also lead to motor action. A probable candidate for what gives rise to the awareness of agency is a copy of the efferent signal that goes out to muscles in order to move. But this connection between efferent signals and the sense of agency prejudices the question of causal involvement of awareness. If we take the sense of agency to be the primary way in which

awareness is involved in action, then we find that it always follows motor processes, never causing them. This is no surprise. We have already eliminated any possible causal role for awareness by making it a common *effect*: one process leads both to movement and to the sense of agency. If we then absorb the other avenues by which conscious awareness could be causally involved in action into the sense of agency, as Gallagher does, we are led to the conclusion that awareness has no substantive causal role in any given movement. But we are not led to this conclusion by the scientific results themselves; we are led to it by the scientific results conjoined with an exclusive focus on the sense of agency, and a failure to sufficiently consider other possible aspects of conscious involvement in action.

If research on the sense of agency is taken to illuminate the causal role of conscious agency, the game is already up before the experiment begins. The premotor processes leading to action will always be a necessary condition for awareness of agency, which means that by the time we are aware of agency, the neuronal preparations for movement are already in full swing. It precludes awareness of agency from occurring before these processes are engaged. Once again, we see that if this is the only potential causal role for conscious awareness, then the reasonable conclusion to draw is that there is no such causal efficacy. But there is no justification for the philosophical moves that make it seem as if this is the only potential causal role for conscious awareness to play.

In sum: Gallagher assimilates other aspects of awareness, including perceptual awareness of the details of what one is doing, to the conscious sense of agency. Other forms of awareness feed into the conscious sense of agency by providing mechanisms by which we monitor our actions and compare them against intentions and goals, so as to make corrections when needed. But this ultimately ignores the causal roles played by these other aspects of conscious awareness insofar as they are *not* solely giving rise to the sense of agency. Gallagher's (2007) goal is to review the scientific study of agency and what it contributes to our understanding of agency from philosophy. By ignoring other forms of conscious involvement and considering this question only with regards to the sense of agency, however, will lead us to think that awareness has at best a subsidiary involvement in movement.

Section 2

Conscious awareness and causality are both commonly described in terms of a certain kind of *directionality*. This shared feature makes it easy to conflate the two, yet the directional relation in the two cases is markedly different. In order to get at the question of whether and in what ways awareness is causally influential in action, the ‘arrow’ of intentionality of awareness must be distinguished from the ‘arrow’ of causal influence. Both of these arrows must be kept track of, but kept track of separately.

This point is straightforward but deserves to be carefully put. An intrinsic feature of awareness is that it is awareness *of* something. Awareness has content and by dint of that content, can be thought of as directed at something, its object.² The object of awareness, that at which awareness is directed, can vary greatly. We are capable of being aware of, having awareness directed at, features of the world: I am aware of the coffee cup on the table. It can be directed at features of the body: I am aware of being hungry. Awareness can be introspective, by having as its object features of experience itself: I am aware of feeling distracted by the noise outside.

The basic problem with the dominant approach is that it utilizes awareness as it is directed at one’s own agency. This is an introspective awareness – the object of awareness is the agency in question. This is true whether we think of the sense of agency as a primitive feature of experience, or as a higher-order reflective awareness. Because when we pose the general question of whether, and if so how, conscious awareness is causally efficacious in our actions, there is a *separate* element of directionality invoked, namely, that of causation. There is a tempting analogy between the arrow of direction of awareness and the arrow of causal influence. Unless one explicitly considers these two distinct modes of directionality, it is easy to simply assume they are the same: that the direction of awareness is the same as that of causal influence; or that, even if the two arrows are not identical, they must point in the same direction. In combination with a focus on the sense or awareness of agency, we get the odd result that awareness must somehow act on separate agentive processes if it is to act on anything at all.

² This notion of intentionality was introduced by Brentano, and is quite widespread and basic, but it bears repeating so as to highlight the contrast with causation.

The central picture is this: if we focus on the awareness of agency and we think that the causal influence of awareness is directed at the same object as is the awareness itself, then what we get is a navel-gazing causal picture. The awareness of agency is introspective. When this introspective view of conscious agency is combined with the conflation of the object of awareness and causal influence, we get what I call the Micromanagement Model of conscious agency: if awareness is to be causally efficacious in action, it must somehow reach down into the bowels of the brain and fiddle with the knobs. In order to have a causal effect on action, awareness must be directed at and then change some otherwise ongoing, unconscious process, perhaps a brute physical process, that is in ourselves and that is leading to the action.

Put this way, it is an obviously silly view: why should awareness have to be directed at our own brain processes in order to make a causal difference? That is why the moniker is appropriate – it assumes that our conscious selves would have to micromanage the details of how our total selves are going about the business of moving around. But, while this way of putting it highlights the absurdity of such a view, this is what we naturally fall into thinking by failing to distinguish the two relevant directionalities and then focusing on the awareness of agency. This double-move separates our awareness of agency from the agency itself, and then turns awareness inwards to monitor our own agentic processes, so that awareness must somehow influence our own agency to be causally efficacious.

Let's see exactly what goes into this view. The Micromanagement Model of conscious agency, the generalized view of conscious agency that emerges from the focus on the awareness of agency, can be characterized as follows: if conscious awareness has a causal influence on action, and the relevant aspect of awareness is that which is directed at our own agency, then any causal influence of awareness in agency will be *on* agency. Our awareness is directed at and acts on our own agentic processes.

The MM has a fairly immediate epistemic consequence for how we should investigate conscious causal influence. If awareness is directed at the processes leading to action, then we should be aware of the elements of those processes. The contrapositive also follows: insofar as we fail to be aware of some causal factor involved in action, then awareness must not have been directed at those processes. That awareness failed to be

directed at some element of the relevant processes means, under the MM model, that awareness was not causally involved in the action that resulted from those processes. Thus, it becomes a necessary condition for awareness having been causally involved that we be aware of the separate processes leading to action. This use of introspection to get at the mechanisms of action is the Epistemic Consequence of MM: if there is conscious causal influence in action, then awareness should be able to provide introspective reports on the causal influences on action, because it is introspectively directed at the processes by which such actions occur.

The Micromanagement Model allows for the possibility of awareness directed at agency without thereby also being causally involved in agency; awareness of agency is not a sufficient condition for causal involvement. But it does make introspective awareness a necessary condition for causal influence: when awareness of agency fails, because we lack introspective access to some causal factor involved in the exercise of agency, this indicates a failure of conscious causal involvement in action. On this view, we cannot be causally efficacious conscious agents unless we are *also* able to provide introspective reports of how we wield that influence.

This assumption of a need for accurate introspective reports of the mechanisms of agency in order for conscious awareness to count as causally efficacious in action can be found in Gazzaniga (1998, 2000), Wegner (2002), Nisbett and Wilson (1977), and in most of Benjamin Libet's work, as well as that of authors who use his work to establish limitations on the causal efficacy of awareness in action (Haggard and Eimer 1999). What these positions share is the assumption that conscious agency is dependent on introspective access. Insofar as we lack conscious access to the mechanisms by which we act, or to factors that influence our action, or to the nonconscious processes that lead to action, we are causally inefficacious as conscious agents.

The separation of agency from awareness of agency is a crucial condition for the Micromanagement Model. Under this view, awareness, whether or not it is causally efficacious in action, is outside of the flow of an otherwise self-sufficient process leading to action. The implicit idea seems to be that there are neuronal processes which lead up to movement; every action involves some kind of movement; and so in order for conscious awareness to be causally involved in producing an action, it must be causally influential

on those processes which are already leading up to the movement which constitutes the action. This leaves the possible effects of awareness as directed at and changing features of these otherwise sufficient processes leading to movement. This is why I call it Micromanagement: nonconscious processes are already in gear leading up to action, and if awareness is to be causally involved, its influence must be on precisely those processes already leading to action. Awareness must reach down into the brain, as it were, and poke at processes already in motion. If we start by thinking that awareness is directed at our own agency, it is a short slide to the requirement that it be causally efficacious by acting on that agency at which it is already directed. Once we accept something like this requirement, we naturally assume that awareness must have access to the causally relevant factors leading to action, since it is already directed at those processes (is awareness of that agency). And this provides a convenient but, I claim, mistaken way to establish whether or not conscious awareness is involved in any given action: by seeing if we can produce introspective reports of our awareness of all the factors that influenced some behavior.

This overall MM view and its epistemic consequence manifest in a variety of philosophical approaches to conscious agency. Under MM, the possible stances on the causal involvement of conscious awareness in agency are quite limited: leaving awareness out of the loop entirely, so we are unaware of what is happening and not causally involved in it (see Gazzaniga 1998); construing awareness as able to report on these processes without affecting them, (see Wegner 2002); or treating awareness as reporting on processes while also interfering with them (thus micromanaging the processes already leading to action). This latter view is not advocated as such by any author, and for a very good reason: it is extremely improbable that conscious awareness somehow reaches down into the brain and fiddles with lower level processes in order to exert causal influence. The point, however, is that once we start thinking of conscious agency in terms of awareness of agency, the last option is the only one which leaves awareness with any genuinely efficacious role in action. But it is only because of implicit acceptance of something like the Micromanagement Model that this strange option is the only one left where conscious awareness has causal efficacy in action. If we assume

something like the MM, the only sensible thing to do is to deny a causal role to conscious awareness.

Conclusion

Gallagher is only one example of a broad trend in contemporary philosophy of psychology that addresses conscious agency by focusing on the sense or awareness *of* agency. This has the effect of separating awareness from agency so that awareness can then be turned inwards and directed at our own agency. This focus on the sense of agency is sometimes combined with the recognition that awareness is involved in action in other ways than as sense of acting, but too often these other avenues of involvement are considered only in terms of what they contribute to the sense of agency.

I have argued that the focus on awareness of agency as the relevant target for investigating the causal efficacy of conscious agency is misplaced. This dominant focus can be generalized as the Micromanagement Model, in which conscious awareness is assumed to be introspectively or internally directed, to reach down into the brain and alter ongoing neuronal processes, and, as an epistemic consequence, to be able to provide a report of those processes in order to be counted as genuinely causal. When we fail to find these features of conscious influence in experimental results, many authors have claimed that this failure provides empirical justification for the position that conscious agency is illusory, inefficacious, or generally confabulatory. I have shown that these conclusions do not follow from the evidence itself, but instead from an assumption about how conscious awareness could be involved, namely as awareness of agency. The Micromanagement Model makes sense of both the tendency to focus on the awareness or sense of agency, and the importance placed on instances of failures of introspective awareness regarding factors that are known to be influential on behavior, both of which are commonly seen in current discussions on conscious agency. It also clearly illustrates why we should take care to avoid this way of thinking about the causal contributions of awareness to agency.

References

- Bayne, Tim and Levy, Neil (2006), "The Feeling of Doing: Deconstructing the Phenomenology of Agency", in Sebanz and Prinz (eds.) *Disorders of Volition*. MIT Press.
- Carruthers, Peter (2007), "The illusion of conscious will", *Synthese* 96: 197-213.
- Chalmers, David (1996), *The Conscious Mind*. Oxford University Press.
- Choudhury, Suparna & Blakemore, Sarah-Jayne (2006), "Intentions, actions, and the self", in Susan Pockett, William P. Banks & Shaun Gallagher (eds.), *Does Consciousness Cause Behavior?* MIT Press.
- Cole, Jonathan (2007), "The phenomenology of agency and intention in the face of paralysis and insentience", *Phenomenology and the Cognitive Sciences* 6: 309-325.
- Gallagher, Shaun (2000), "Self-Reference and Schizophrenia: A Cognitive Model of Immunity to Error through Misidentification", in Zahavi (ed.) *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-experience*. Amsterdam and Philadelphia: John Benjamins, 203-239.
- Gallagher, Shaun (2007), "The Natural Philosophy of Agency", *Philosophy Compass* 2: 347-357.
- Gazzaniga, Michael (1998), *The Mind's Past*. Berkeley, CA: University of California Press.
- Gazzaniga, Michael (2000), "Cerebral specialization and interhemispheric communication: Does the corpus callosum enable the human condition?", *Brain* 123: 1293-1326.
- Haggard, Patrick (2003), "Conscious Awareness of Intention and of Action", in J. Roessler and N. Eilan (eds.), *Agency and Self-Awareness*. Oxford University Press, 111-127.
- Haggard, Patrick and Eimer, Martin (1999), "On the relation between brain potentials and the awareness of voluntary movements", *Exp Brain Res* 126: 128-133.
- Horgan, Terence E.; Tienson, John L. & Graham, George (2003), "The Phenomenology of First Person Agency", in S. Walter and H. D. Heckmann (eds.), *Physicalism and Mental Causation: The Metaphysics of Mind and Action*. Imprint Academic, 323-40.
- Jack, Anthony and Shallice, Tim (2001), "Introspective physicalism as an approach to the science of consciousness", *Cognition* 79: 161-196
- Libet, Benjamin (1985), "Unconscious cerebral initiative and the role of conscious will in voluntary action", *Behavioral and Brain Sciences* 8: 529-566.
- Metzinger, Thomas (2006), "Conscious volition and mental representation: Toward a more fine-grained analysis", in Natalie Sebanz & Wolfgang Prinz (eds.), *Disorders of Volition*. MIT Press.

- Mossel, Benjamin (2005), "Action, Control. And Sensations of Acting", *Philosophical Studies* 124: 129-180.
- Nahmias, Eddy (2005), "Agency, authorship, and illusion", *Consciousness and Cognition* 14: 771-785.
- Newell, A. (1992), "SOAR as a unified theory of cognition: Issues and explanations", *Behavioral and Brain Sciences* 15: 464 – 492.
- Nisbett, Richard and Wilson, Timothy (1977), "Telling More Than We Can Know: Verbal Reports on Mental Processes", *Psychological Review* 84: 231-259.
- O'Shaughnessy, Brian (2003), "The Epistemology of Physical Action", in J. Roessler and N. Eilan (eds.), *Agency and Self-Awareness*. Oxford University Press, 345-357.
- Velmans, Max (2004), "Why conscious free will both is and isn't an illusion", *Behavioral and Brain Sciences* 27: 677.
- Wegner, Daniel (2002), *The Illusion of Conscious Will*. MIT Press.
- Wegner, Daniel M. & Wheatley, T. (1999), "Apparent Mental Causation: Sources of the experience of will", *American Psychologist* 54: 480-492.