

Conceptual Constraints in Empirical Investigations into Consciousness and Behaviour

9406903

Msc Philosophy

The University of Edinburgh

2009

I have read and understood The University of Edinburgh guidelines on Plagiarism and declare that this written dissertation is all my own work except where I indicate otherwise by proper use of quotes and references.

Contents

<u>SECTION 1 - CONCEPTUAL CONSTRAINTS: A BRIEF PROLEGOMENA</u>	<u>2</u>
1.1 - THE COLLISION OF SCIENCE AND POPULAR NOTIONS OF CONSCIOUSNESS.....	2
<u>SECTION 2 - DANIEL WEGNER AND THE ILLUSION OF CONSCIOUS WILL</u>	<u>5</u>
2.1 - THE CONSCIOUS WILL: A CONCEPT	8
WHAT IS THE CONSCIOUS WILL ANYWAY?	8
THE DIRECT EXPERIENCE OF MENTAL CAUSATION	9
A CONSISTENTLY PRESENT EXPERIENCE?	11
2.2 - THE ILLUSION.....	13
AUTOMATISMS: OR THINGS YOU DO WITHOUT NOTICING	14
I SPY	17
2.3 - THE ILLUSION: A CRITICAL RESPONSE	20
WEIGHING THE EVIDENCE	21
CRITIQUING THE STRUCTURE	22
PETER CARRUTHERS: A DEFENSE.....	24
BENJAMIN LIBET	26
BRAIN BEFORE MIND	27
GALLAGHER AND APPROPRIATE TIMESCALES	30
<u>SECTION 3 - THE PHENOMENOLOGY OF AGENCY</u>	<u>33</u>
3.1 – THE IMPORTANCE OF PHENOMENOLOGY FOR CONSCIOUSNESS STUDIES	35
PHENOMENOLOGY MATTERS	36
3.2 – AGENTIVE SELF-KNOWLEDGE: EXPERIENCE VS. JUDGMENTS	37
3.3 - SOURCES OF OUR EXPERIENCES OF AGENCY	39
3.4 – AGENCY IS A PROPERTY OF AGENTS NOT MECHANISMS.....	41
<u>CONCLUDING REMARKS.....</u>	<u>43</u>
<u>REFERENCES</u>	<u>44</u>

Section 1 - Conceptual constraints: a brief prolegomena

This paper is based on the very simple idea that when we begin scientific investigations of any kind we bring to bear conceptual assumptions that need to be accounted for. I call this idea the conceptual constraints of empirical investigations, and in its simplest and most straightforward form it is entirely uncontroversial. Conceptual constraints are a part of every investigation, but with varying degrees of clarity and precision, that is, there are times when greater effort is put into defining exactly what is being probed and why this object is the relevant one for a certain question. This constraining is an important part of our research because the conclusions of our scientific explorations depend upon our experimental paradigms actually probing the ‘thing’ that we are making claims about; accounting for our assumptions gives greater validity to our conclusions. So what considerations are important to factor into our experimental targets in consciousness studies? How do we ensure that we do not make conclusions about conceptual straw men, that is, conclusions about nothing? This problem seems to be especially hard for consciousness studies as it is a very interdisciplinary field, and interdisciplinary discussions always seems to exacerbate methodological and conceptual differences on any given problem (See Maasen 2003, 303), and because when it comes to consciousness everyone takes herself to be her own best guide and expert, for we all think we know with some degree of authority the nature of our own conscious experience. However, despite these challenges, and in fact because of them, conceptual constraints are that much more important for investigations into consciousness so as to give our conclusions wide ranging appeal and impact.

1.1 - The collision of science and popular notions of consciousness

The current situation in consciousness studies is one where popular notions of the mind are consistently under attack from neuroscientific research, and of course this is really only a new twist in the ongoing scientific correction of folk psychology going back to James and beyond. Sellars’ (1962) picture of the clash of scientific and manifest images is perhaps appropriate here, as it is the case that some researchers perceive themselves as supplying the evidence that undermines the dubiousness of the manifest image of consciousness. Daniel Wegner is one of these researchers and his claim that the conscious will is an illusion is his effort to liberate us from the seductive intuitions that

lead us astray. “The fact is, we find it enormously seductive to think of ourselves as having minds, and so we are drawn into an intuitive appreciation of the conscious will (Wegner 2002, 26).” This sort of claim is a perfect example of the kind of conclusion, made on the basis of Wegner’s extensive empirical research, that requires a robust explanation of its conceptual targets in order for it to be convincing. It may be the case that our belief that we have minds is based on its seductive nature, but to prove this requires an appropriately rich and relevant account of what it is we actually believe, that is, what our intuitive concept of the conscious will contains. The problem we can see in Wegner’s work, however, is that too little effort is put into rigorously defining the concept of conscious will. Bertram Malle is someone else who is critical of this type of scenario, and he puts the problem with Wegner and others like him this way:

Are they going to talk about these things the way ordinary people do (that is, are they going to use the folk concepts with their normal meaning and reference), or are they introducing technical terms? ... Ultimately, whenever a scientist describes a folk assumption and claims that some evidence shows it to be wrong, the scientist has to adhere to the folk meaning of the relevant concepts, or else the evidence isn’t relevant (Malle 2006, 208-209).

The call for conceptual constraints simply represents the need for general conclusions like Wegner’s to reflect either the general understanding of a concept or to argue for a special sense of a term. Even if there is a debate over what should be a constraining factor for a concept, for example whether or not phenomenology is valid for informing our concepts of conscious mental states, it is still reasonable to expect Wegner, or anyone else, to show how his conclusions are conclusions about what we actually experience.

The in-depth critique of Wegner’s work on the conscious will in this paper will make two conclusions obvious. The first one is that even though it is important for science to provide constraints to our theoretical problems (See Ross 2006, 126), the opposite is also true, and Wegner does not offer us a convincing argument about the illusions of our agency because he does not take measures to incorporate appropriate conceptual constraints. The second conclusion arises from our observations of Wegner’s failings, and it is that we can improve our empirical efforts by employing phenomenological insights into our conceptual frameworks. Much of the weakness in Wegner’s account is produced by his lack of phenomenological considerations, and this suggests that a way

forward for our ongoing empirical research into consciousness and behaviour should involve our conceptual targets being formed with phenomenological insight.

Section 2 - Daniel Wegner and the Illusion of Conscious Will

“All theory is against the freedom of the will; all experience is for it.” Samuel Johnson – Boswell’s *Life of Johnson* (1791)

If there is one idea that succinctly expresses Wegner’s position on the will, it is this one from Samuel Johnson. Wegner’s goal, in his many articles and books on the subject of conscious will, is to show that despite the sense we get from our experience, the will is not what we think it is; all ‘theory’, that is, empirical evidence, is against the folk intuition of the power of the conscious will. So, according to Wegner, we are caught in a tension between experience and evidence, and as a well-informed psychologist he understands the balance to be tipped in favor of ‘theory.’ Wegner places himself in the company of William James, who puts it this way “The whole sting and excitement of our voluntary life...depends on our sense that in it things are *really being decided* from one moment to another, and that it is not the dull rattling off of a chain that was forged innumerable ages ago (James 1890, 453).” Thus in all of his research, Wegner’s target is the elucidation of the “dull chain” such that the sense of “things really being decided” is shown to be what it is, an illusion. He says,

The fact is, it seems to each of us that we have conscious will. It seems that we have selves. It seems we have minds. It seems we are agents. It seems we cause what we do. Although it is sobering and ultimately accurate to call all this an illusion, it is a mistake to conclude that the illusory is trivial. On the contrary, the illusions piled atop apparent mental causation are the building blocks of human psychology and social life (2002, 341-342).

If nothing else, hearing Wegner’s conclusion of “apparent mental causation” in such stark terms should impress upon us the seriousness and scale of his project, which is nothing less than the liberation of all humanity from the foolish and unfounded beliefs it uses to function daily. For this is what he is claiming: everyone, everywhere moves along according to some inescapable causal chain and no one has any idea that this what is actually going on.

Taking this claim seriously, as we should, requires that we expect significant evidence to be forthcoming. Wegner believes he has this evidence and indeed, the material Wegner marshals to explain and defend his claim is impressive. All sorts of research on the timing of neurological processes, the accuracy of introspection, and the experiences of

hypnotism and automatism is presented in order to free us from the self-deceiving state in which we are located. However, and this is a crucial point, the evidence is misdirected; he targets his research at something called ‘the conscious will’ without constructing an appropriate account of what the conscious will is. We are, apparently, just supposed to all know, and agree on, what it is, and this assumption is a problem. Without a doubt Wegner identifies unique experiences of agency and behavior that require explanation, and any competing account of the causal mechanisms of human behavior needs to make sense of the phenomena he identifies, but without first showing that the concept of conscious will that he is targeting is either the concept that most people have, or that he has good reason for formulating the concept in a particular way, as a technical term, he constructs an argument to show how the conscious will is an illusion. In fact, there is good reason (as I will discuss below) to understand the relationship between the conscious mind and behavior in ways other than Wegner assumes, that is, there are other conceptual models of how we might know about and cause our behavior that do not fall prey to the challenges Wegner presents. What I will show through an in-depth exploration of Wegner’s research, his experiments and their analysis, is that his conclusion of illusion depends on this specific and ill-formed concept of the conscious will, specifically, his assumption that the conscious will is generally taken to be: (1) an experience of mental causation, that is, a direct experience of conscious thoughts (intentions) causing an observed action, (2) that the representational content of the conscious will is immune to error, and (3) that all voluntary action prompts an experience of the conscious will.

This section of the paper is in three parts. In section 2.1 I comb through Wegner’s work to give an outline of his conceptual model of the conscious will. In 2.2 I examine the experimental evidence he offers to show how the conscious will as a experience of mental causation is an illusion and that the true nature of the conscious will is that of an inferential mind-reading system. Finally in 2.3 I present some criticisms of this evidence and its defenders in order to demonstrate how Wegner’s argument fails and fails because he targets the wrong thing. After this, the following and final sections of the paper 3.1-3.3 presents material that support an alternative model of consciousness and agency to both to the model Wegner attempts to discredit and to the solution he offers. This

alternative model does justice both to many of our intuitive notions of consciousness and also to the psychological and neurological evidence that Wegner claims for himself, while even incorporating some of the lessons learned in Wegner's observations. Not only does Wegner's argument of illusion attack a straw man, so to speak, but the counter explanation he offers to correct this false image fails to satisfy the phenomenal qualities of our experience of being agents.

2.1 - The Conscious will: a concept

Wegner's comments on the concept of conscious will fall into two groups. There are those comments that he makes in an attempt to show us what it *really* is, and there are those comments that are trying to explain what it is that we *take it to be*. The critical reading of Wegner's work in this section is mostly trying to clarify this second thing, because in order to appreciate how the conscious will is an illusion we need to know exactly what the experience is that is an illusion. Hopefully without being confusing this will sometimes be done by comparing the purported illusory experience with Wegner's 'fact-of-the-matter' alternative.

The goal in this section is simply to explore how Wegner employs the term conscious will as a target for experimental evidence, and this exploration unfolds in roughly two stages. The first of these stages comments on the most obvious observation of Wegner's account, which is that the conscious will is just plain poorly defined. The way in which Wegner explains the experience of the conscious will and its representational content is confusing and sometimes contradictory. The second half of this exploration looks at the many strands relating to the central element of the conscious will, which is the idea that it is a "theory of behavior causation (2002, 325), or "apparent mental causation (1999)." The combination of the two images of Wegner's conceptual model exposed in these stages incorporates the large majority of what he says regarding the conscious will and sets the stage for his argument that the conscious will is an illusion.

What is the conscious will anyway?

The criticism that Wegner presents us with a poorly developed model of the conscious will is certainly not limited to the present work. Tim Bayne is a critic who also sees Wegner as employing a poorly developed model of the experience of being an agent and that this "naïve model" is at the heart of his "will-skepticism (Bayne & Levy 2006)." Bayne identifies in Wegner's work several descriptions of the content of the conscious will that taken together leave us with a challenge for understanding the exact merits of Wegner's claims. He notes,

For the most part Wegner leaves the notion of the experience of conscious will at an intuitive level, often describing it simply as the 'experience of doing.' He does, however, link the experience of conscious will with various other aspects of the

phenomenology of agency, such as the experience of authorship (Wegner 2005, 27), the experience of *intentionality* (Wegner and Erskine 2003b, 688), the experience of *effort* (ICW, 39), the experience of *free will* (ICW, 318), and the experience of *mental causation* (Wegner 2005, 23) (Bayne 2006, 170).

This exegesis provides a useful entry point into Wegner's work because the experiences identified by Bayne cover the bulk of references Wegner makes to the conscious will.

One of the first things to notice in considering these experiences as they relate to the conscious will is that they are far from being synonyms, but apparently are used as such. The experience of doing does not necessarily contain the phenomenal quality contained in the experience of effort, something like the quality of energy or ego depletion. Doing something may sometimes, or even frequently, be effortful, but our ability to distinguish between the two qualities demonstrates their distinctiveness, however slight. Likewise, the experience of effort is certainly not the same thing as intentionality or mental causation. Intentionality, as Wegner is using it, is a concept of goal or purpose oriented action, and effort carries no such quality. And neither intentionality nor effort is equivalent to mental causation. The experience of mental causation is presumably an experience of some behavior being caused by a mental cause as opposed to some other kind of cause. The upshot of all this for the notion of the conscious will is that it is an experience that represents a behavior as not merely caused by an agent, but caused by an agent *in a certain way*, that is, the conscious will is an experience of the activity of mental causes. Thus if we take the conscious will to be an experience that incorporates all these elements we must either see it as being internally inconsistent or as being a broad catch-all kind of concept that contains a range of overlapping and distinct experiential qualities. It is probably fair to assume that Wegner takes it to be something like the latter.

The direct experience of mental causation

Aside from the generally confusing discussion of the nature of the conscious will, there are a few key ideas that Wegner seems to focus on and are particularly important for his argument. It is to these that we will now turn. Wegner argues that we are under the illusion that our experience of conscious will informs us directly about the sources of our behaviour, that is, our conscious thoughts. This is an illusion because the mechanisms producing the experience, according to Wegner, are not 'direct' but are in fact

interpretive and inferential. Wegner's argument of illusion, therefore, attempts to prove that the source of our experience is interpretive but it also requires that our experience is otherwise. Thus we can see both sides of Wegner's use of the conscious will, what he claims the conscious will *actually* is and what it is *not*, by looking at some of his points side by side. The picture of an interpretive and inferential conscious will is perhaps best seen in the comparison Wegner makes to the relationship between a ship's wheel and its compass.

Conscious will is the mind's compass. As we have seen, the experience of consciously willing action occurs as the result of an interpretive system, a course sensing mechanism that examines the relations between our thoughts and actions and responds with "I willed this" when the two correspond appropriately. This experience thus serves as a kind of compass, alerting the conscious mind when actions occur that are likely to be the result of one's own agency (2002, 317).

The important thing to see in this comparison is actually the dissimilar elements between a ship's compass and "the compass of the mind." The difference is that in this quotation we can see that Wegner portrays the conscious will as being related to actions as a measure of *likelihood*, not a relation of regular and consistent mechanism as in the case of the ship and its compass, which is bound physically to the vessel and to planetary magnetism. This is not necessarily an inappropriate sense of how we might understand our self-knowledge with regards to our own agency, but it does provide a basis for understanding the next two, and most important, aspects of Wegner's conscious will.

Contrasting with the claim that our sense of agency is an experience of interpretation, Wegner stipulates that the conscious will does not present itself as such. The experience of conscious will is not an experience that has a quality of interpretation; we do not notice that we are making judgments. It simply is an experience of mental causation; there is a sense of directly experiencing an action being caused by a mental source. We can see this element of Wegner's account of the conscious will by looking at the two following claims:

Why does it feel as though we are doing things? The experience of consciously willing our actions seems to arise primarily when we believe our thoughts have caused our actions (2005, 23).

“This means though that the conscious will is an illusion. It is an illusion in the sense that the experience of consciously willing an action is not a direct indication of that the conscious thought has caused the action (2002, 2).

By arguing that the illusiveness of the conscious will is the fact that the conscious will does not directly indicate mental causation, Wegner obviously implies that this is what he takes the experience of conscious will to be. It is an experience with a phenomenal quality of being a *direct* experience of conscious thoughts causing action. Another way of putting this sort of quality is that the experience of mental causation is one that we have personally and internally with immunity to error. By experiencing it directly, the implication is that we know, with accuracy only we can have, the nature of the experience.

A consistently present experience?

The final component of the concept of the conscious will fits side by side with its immunity to error, and this is its omnipresent awareness with regards to presence of voluntary action. Omnipresent awareness just means that for Wegner, the conscious will is an experience that is supposed to *always* occur when we perform a voluntary action. The model of conscious will that Wegner targets with his experimental evidence, is that for any voluntary action performed by a conscious human agent, the mechanisms that produce the action are tied to the experience of conscious will so that we could not perform a voluntary action without the appropriate experience. As we will see in the next section, the clearest source of this idea is Wegner discussion of automatism. For Wegner, automatism, behavior where there is no sensation of intentionality, are a phenomenon that requires explanation precisely because of this lack of sensation. Therefore if Wegner takes these unique behaviors as demonstrating the lack of something essential and that this demonstration is a link in his “illusion” argument, then he must understand the concept of the conscious will that he is critiquing as being an experience that is always aware of voluntary action.

Our snapshot view of Wegner’s conceptual model of the conscious will, the experience that he believes to be illusory, shows the conscious will to be: (1) broad and un-refined, containing a multiplicity of experiences that are not necessarily similar enough to be

synonyms, (2) a direct experience of mental causation, (3) being direct it is therefore immune to error, (4) it is always present in voluntary action.

2.2 - The Illusion

“What then is the illusion? The illusion I hoped to explore in *ICW* was the illusion of a *self-knowing mechanism* (Wegner 2004, 682).”

“When a person thinks of diving into a pond and then does so (and feels this was consciously willed), we normally accept that the causal mechanism underlying this event knows itself (2004, 682)!”

Wegner’s claim that the conscious will is an illusion, or “the mind’s best trick,” is his solution to the very simple question: how is it that we can be so wrong about our conscious involvement in the causal chain of an event? We claim to not be involved when we are, as in the case of table turning and other automatisms, and we claim to be involved when we are not, as in the case of Wegner’s well-known experiments ‘I Spy (Wegner & Wheatley 1998)’ and “Helping Hands (2004)”. As Wegner might put it, if our conscious experience plays a real part in how behaviour comes about, then we should be much better at knowing when we did something than we are. He suggests alternatively, “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 (2005, 65).” The illusion, then, is not just that we can be wrong about what we do, but that our experience appears to us to come about in a certain way, as a self-knowing mechanism, and seeing as how the experience systematically represents the world incorrectly, the content of the experience, the fact that our action has a conscious thought as a cause, must be wrong. The experience of conscious will, according to Wegner, is an experience that we take to be a reliable source of information about our agency, but because it is so frequently wrong, it must be the case that our agency is of a different character than our experience suggests. To put a nail in the coffin of the veridicality of this experience, Wegner also argues that the physiological processes that produce the conscious will are not even the same ones that produce our actions, and this separation definitively rules out any conscious involvement in action causation.

This section of the paper examines the evidence in support of Wegner’s argument that we: (1) do not always know when we are causing actions, i.e. the conscious will is not always present when it should be, (2) that it is not infallible because we can actually

manipulate the experience of conscious will according to certain principles, (3) that this ability to manipulate the experience points us towards the true cause of the conscious will - an inferential system of image production. Wegner also claims that the experience of conscious will is physiologically distinct from action producing mechanisms. This is really a separate argument that Wegner puts forward, and is an important claim that cannot be sidestepped, so will be examined here as well but in a following section (2.3). What becomes apparent in examining this supporting material is that there are many interesting experiences related to the sense of agency, and Wegner is right to call attention to them. However, the evidence also raises some problematic issues for Wegner's account. These issues center on the assumption that all voluntary actions require an experience of conscious will and that the experience must have direct access to causal mechanisms.

Automatisms: or things you do without noticing

The first leg of this three-legged argument is Wegner's exploration of the phenomenon known as automatisms. Automatism is a term meant to catch those movements that occur in a certain relation to people but without any feeling from those involved played that they played a causal role. The examples Wegner gives are of the popular 'paranormal' interactions from the nineteenth century that are no longer practiced widely today. These are the experiences of table turning, automatic writing, Ouija boards, etc. The following is Wegner's general comment on these experiences:

The absence of conscious will in automatism, then, is particularly profound. It is more than those little recognitions we have that we seem to have let up on the mental gas pedal. The loss of perceived voluntariness is so remarkable during an automatism that the person may vehemently resist describing the action as consciously or personally caused (2002, 99).

In all of the activities marked as automatism by Wegner, the two defining characteristics are an observed movement being caused by a person and the person in question not reporting any experience of intentionally causing this movement; there is a loss of "perceived voluntariness." In fact, automatisms are, as Wegner explains, remarkable precisely because the movements are obvious, but the source is not because participants deny any involvement. How could this be, Wegner asks, if the conscious will is an experience that faithfully informs us as to instances of mental causation and these actions

are clearly not simple reflexes but patterned behaviour? He gives the example of the Reverend Godfrey who participated in a reported table turning séance in the mid-nineteenth century. Godfrey reports a proliferation of knocking and turning caused by unknown and possibly demonic source, as the table only stopped moving when a Bible was placed on it. Due the widespread popularity of these kinds of events, table-turning was investigated by physicist Michael Faraday where he concluded after rigorous experimentation that the movements were indeed caused by the human participants (2002, 8). Godfrey, it is reported, was shocked and vehemently denied the evidence. For Wegner, the conclusion to draw from this evidence is clear. Godfrey's shock and denial of the evidence only makes sense if the experience of conscious will is assumed by Godfrey to be a reliable and *constant* experience of causation. Godfrey, as well as participants in automatism generally, had no such experience and therefore could not believe that he caused the table to move.

The rest of Wegner's discussion of automatism is comparable to this one, but this story serves to show us the most important point. In using accounts such as this one of Godfrey and Faraday, it is clear that Wegner assumes (1) that all voluntary actions require a corresponding experience of conscious will, and also that (2) these automatism are full-blooded voluntary actions. This is, in fact, something he specifically says, "The experience of will can be reduced to very low levels under certain conditions, even for actions that are voluntary, purposive and complex-and what remains is an automatism (99)." Automatism, like table turning, are 'voluntary and purposive' actions that according to Wegner, do not give an appropriate experiential report, there is, a "loss of perceived voluntariness." Table-turning is evidence of the illusion of conscious will because participants generally believe that they know when they cause a movement (they assume there is an experience of conscious will that tells them!) but, as shown by Faraday, participants in an automatism do not always know when they have acted because they do not always have an experience of conscious will even though they should have. Automatism are simply voluntary actions that fail to have the requisite 'authorship emotion' attached. However, these reports are only evidence for Wegner if you grant that automatism are an example of a conscious, intentional action. Wegner's argument about the conscious will is about how we ascribe self-agency in the presence of

conscious thoughts about an action, so for automatisms to be relevant to his argument they need to be not just complex, purposeful actions that do not require conscious input, but they need to be occur in the presence of conscious thoughts without the experience of conscious will. If automatisms were merely movements that occurred in relation to our body that we did not feel we consciously caused, they would not be any more supportive for Wegner's case than the countless complex, but routine, actions that we engage in daily that never seem to be consciously directed. Automatisms are, for Wegner's argument, a unique situation, but this requires some further explanation that he does not give. Indeed Wegner notices that the lack of conscious thoughts about the particular actions in an automatism is a problem for his account and he attempts to rectify this.

Taking a little longer look at this aspect of Wegner's overall account only further undermines the usefulness of automatisms for overall case. In a quotation given above Wegner mentions that subjects in these scenarios typically deny that the action was "consciously or personally caused." We may grant that in some way these actions are personally caused, as in the mere fact that individual persons are clearly involved in the action, but should we think of them as consciously caused? The whole weird experience seems to be based on the fact that no one involved has any conscious thoughts about the movement, which on Wegner's account of the conscious will being an experience that reports on the presence of conscious thoughts that occur with a certain relationship to an action, this then seems to make sense of the situation and the denial of involvement. No one experiences conscious will, or mental causation, because there were no conscious thoughts causing the actions. If this is the case then some part of the illusiveness of conscious will is put in jeopardy because automatisms cease to be an exceptional circumstance. For Wegner's argument against the conscious will as "apparent mental causation" to work, he needs a subject to have conscious thoughts about an action, and perform an action, but no experience of conscious will, otherwise the lack of subjective reports of conscious will ceases to be a problem.

The stopgap measure Wegner supplies to fill this hole is thought suppression. However, there is some irony to this solution as Wegner himself has done extensive research on the human inability to suppress conscious thoughts.

Can people suppress their conscious intentions? There is a fair amount of evidence that people find it unusually difficult to suppress thoughts, at least for short periods of time in the laboratory. (e.g. Wegner 1989; Wenzlaff and Wegner 2000). So at one level the answer would seem to be no...It may be that people trying to produce automatic writing, Ouija spelling, or similar automatism are sometimes actively drawing their own attention away from the conscious intentions that come to mind, either by direct suppression or by other mental techniques. They may be purposefully avoiding the conscious representation of their purposes, at least for a while until the thought returns. This could then yield a temporary reduction in the experience of the will (2002, 139).

So the final picture of automatism and the conscious will is that according to Wegner the people involved *do* have conscious intentions to cause the action but suppress them thereby not experiencing the conscious will, but because those involved truly are the source of the movement therefore they *should* always feel conscious will and should not be able to side-step it via suppression. If the conscious will is attached to action production it should be immune to these sorts of influences. This sort of picture seems to be weak at best because it both relies on the first-person report that there was no experience of conscious will, but rejects (or explains away) the denial of conscious intentions prior to the action.

The second body of data Wegner supplies to support his argument comes from experiments he has performed himself. These experiments do double duty for his argument because they demonstrate an experience of conscious will that falsely indicates the source of action, and these false indications are produced according to the principles Wegner's claims are the real source of the conscious will. The goal in these studies is to show that conscious will is produced by an inferential, image making system, which, according to Wegner, is not what the experience of conscious will feels like.

I Spy

Wegner's 'I Spy' experiment is one of the paradigms he uses to show how we can be manipulated into believing that we are doing something we are not. The goal of the study is to see whether subjects report a sense of willfully directing a movement when they have not. The experiment consists of a subject who places her hands on a modified computer mouse along with an experimental confederate appearing as another subject. While wearing headphones, the subject moves a cursor jointly with the confederate around a screen that is covered in objects from the children's book "I Spy." At an

apparently, but not actually, random point in time the subject hears a combination of music and words begin in her headphones and she is supposed to point the cursor towards an object on the screen within the next few seconds. After the pointing is complete the subject rates on a seven point scale her intentional involvement in the action, as opposed to the confederate. The experimental construct has the confederate occasionally force a particular choice in close conjunction with the subject hearing the selected object named in her headphones. In these instances where the subject hears the targeted object closely before the action, she typically rates her involvement higher than when the movement is forced without an accompanying audible prime. The audible prime was given at different time intervals before and even slightly after the forced movement, and the influence of the prime decreases with temporal distance from the act. Also, as a control, the experiment tested the influence of audible primes without any direct action by the confederate to see whether the prime alone could cause the targeted movement, which it did not.

There are two conclusions Wegner draws from this experiment. The first is that obvious, for Wegner, if the conscious will was a veridical experience of action causation it should not be open to this kind of influence. This is, of course, a conclusion that (as we have seen) depends on the conscious will being understood as a direct and therefore infallible experience of mental causation. The other conclusion that Wegner makes here is that this kind of manipulation of the experience of conscious will shows the true mechanisms of the experience of conscious will. He concludes,

The experience of willing an act arises from interpreting one's thoughts as the cause of this act. Conscious will is thus experienced as a function of the priority, consistency, and exclusivity of the thought about the action. The thought must occur before the action, be consistent with the action, and not be accompanied by other causes. An experiment illustrating the role of priority found that people can arrive at the mistaken belief that they have intentionally caused an action that they were forced to perform when they are simply led to think about the action just before its occurrence (Wegner & Wheatley 1999, 480).

A more thoroughgoing critique of this conclusion is supplied below, but it is helpful for now just to recall the thrust of Wegner's argument that produces this conclusion. The illusion of the conscious will is that we experience it to be something other than it actually is. Wegner argues that the presumed nature of the conscious will is such that it is

a “self-knowing mechanism”, that is, we experience the conscious will and it is an experience of mental causation, however, the truth exposed by automatisms and “I Spy” is that the experience of conscious will is merely an image of mental causation produced by an inferential system that functions according to the principles of “priority, consistency, and exclusivity of the thought about the action.”

2.3 - The Illusion: A critical response

There are two broad directions that a response to Wegner could take and I attempt to do both, one in this section and one in the final section (3.1-3.3). The first kind of response is to challenge the strength of his argument, which is the focus of this section, and the second response, which is the focus of the final section, is to challenge the validity of his target. Both of these challenges point to the same underlying issue in Wegner's work. There are many experiences related to our sense of agency, but the experience Wegner calls an illusion (the conscious will) does not seem to represent all of the ways we experience our agency. As such, the argument he gives us to show that our experience of conscious will is an illusion does not do much to show that we can know nothing about our agency via our experiences of agency, which is the dramatic conclusion that he makes (i.e. "It *seems* we are agents. It *seems* we cause what we do.")*[italics mine]*. The point is that because Wegner is making a claim about the nature and sources of an experience, it is necessary for him to do the job of clearly articulating the conceptual boundaries of this experience and separating it from other nearby experiences. This section makes clear that even as his empirical results speak to many interesting experiences related to our agency, Wegner's argument does not do justice to them mostly because he does not do the required phenomenology.

Having seen the structure and content of Wegner's illusion argument in sections 2.1 and 2.2, we can now ask (1) Does Wegner's supporting evidence stand up to scrutiny? What do studies on the experience of conscious will show us about our sense of agency? We can also ask (2) Does Wegner's conclusion follow from his premises? Or put even better, even if we grant that we sometimes make inferential judgments about our agentive activity does this mean that we have no experiences that veridically inform our sense of agency? The answer to both these questions will be clear. Neither does Wegner's key evidence provide much needed support nor does his conclusion follow from his argument and therefore illusion is left unproven. This section has three parts with the first addressing question (1) and the validity of Wegner's evidence. The second part addresses question (2) as well as a defense of Wegner's position by Peter Carruthers. In the third part I raise and refute Wegner's other argument for illusion mentioned in section 2.1, which contends there are separate physiological structures for action and experience.

Weighing the evidence

The “I Spy” study is one of the most important parts of Wegner’s arguments because it shows both that the experience of conscious will can occur when it should not, and it shows how this false experience can be triggered. The false trigger is, of course, Wegner’s conditions of priority, consistency, and exclusivity. There is good reason to think, however that both of these parts of “I Spy” are not demonstrated as conclusively as Wegner would like. The first problem with “I Spy” is that Wegner gives the impression that subjects who were primed according to his ‘inferential principles’ responded by having a full-blooded experience of conscious will. He concludes that “the experience of will can be created by the manipulation of thought and action in accord with the principle of priority (2002, 78).” But this conclusion glosses a few important details. It is true that when a subject heard the priming word at the appropriate time, satisfying the conditions of priority and consistency, her sense involvement in the action was heightened, but only slightly so! The average score ranking an involvement in a forced movement rose to 60% from 45%, which clearly says that judgment was influenced but certainly does not show that satisfying these conditions can produce a full-fledged experience as implied by Wegner.¹ A rise to 60% shows that something is still lacking in the experimental circumstances such that a full experience does not occur. If we can only be slightly influenced, and this is even under complicated experimental circumstances, to falsely have a sense of involvement in an action, then it seems possible we do have non-inferential kinds of knowledge about our agentic behavior.

To put another crack in this particular bit of research, neurological research on patients with schizophrenia has shown again that Wegner’s conditions of priority, consistency, and exclusivity are not sufficient to produce an experience of conscious will, or an experience of conscious intentions causing behavior. In the peer commentary to Wegner’s précis of *The Illusion of Conscious Will* a team of researchers including Chris Frith, Johannes Schultz, and Natalie Sebanz report on their investigations into schizophrenic patients and the lack of agentic experience that occurs in this pathology. They state,

¹ Cf. Nahmias 2005.

Patients with delusions of control report that their actions, even quite trivial actions, are being controlled, not by themselves, but by some alien force. Patients report such abnormal experiences even though they have the prior intention to make the action, the action made is consistent with their intention, and there is no obvious ambiguity about who is making the action. We have suggested elsewhere (Hohwy & Frith 2004) that what is missing is an aspect of the feeling of what it is like to be in control of one's actions; knowing what is going to happen and at the same time minimal awareness of the sensory consequences (In Wegner 2004, 674).

This kind of evidence augments the previous criticism of 'I Spy', in that again the conditions given by Wegner are just not sufficient to produce the experience of being an agent. Frith et al. also point to a sensory source, most likely proprioceptive, that is needed to give a person a full-blooded experience of consciously willing an action. We will return to pathological experiences of agency, like schizophrenia in the final section because as indicated in this quotation, there are important phenomenal qualities of being an agent that are apparently pathologically dissociable and point us towards a more complex model of the sense of agency than the one Wegner is challenging.

Critiquing the structure

Moving on the structure of Wegner's argument we can see that the crucial claim that because the source of our experience of conscious will is apparently inferential we must disregard its convincing nature because our experience *cannot tell us reliably about our agent causal activity*. My doubts about Wegner's take on the nature of our "convincing experience" will be developed below, but for now the most important question is why interpretation or inference should prove either unreliability or illusion. Here is a brief recap of Wegner on Wegner:

We only understand mental causation in ourselves by virtue of an authorship processing system that examines a variety of indicators to determine whether this particular action was the one that we willed.

Ultimately the illusion of conscious will comes down to an issue of the validity of self reports (2004, 682).

If we assume for the sake of argument that the first of these statements is true, then we can also agree that the second is true, however, it does not seem obvious that this second claim actually works in Wegner's favor. Eddy Nahmias identifies the problem with this point.

Suppose the theory of apparent mental causation is accurate in that our experience of ourselves or our intentions as the cause of our actions is inferential. The fact that we experience such causation as *non*-inferential does not show the experience to be illusory unless the conclusions of the inference is systematically mistaken. So it would have to be shown not merely that we do not know everything about the process by which we experience our actions as caused by our conscious intentions, but that, in general, the intentions (or other relevant mental states) that we are (sometimes) conscious of are *not* the cause of our actions (Nahmias 2005, 777-778).

It seems that in order for our experience, even as Wegner has portrayed it, to be an illusion it would have to be the case not just that we have a fallible system producing the conscious will but that our conscious thoughts *do not* cause our actions. Even if we do not directly experience mental causation, the validity of an authorship system indicating mental causation is given by whether or not mental causation exists, not by whether or not we experience it.

In a way that only makes things more confusing and does nothing to advance his argument, Wegner sometimes suggests that the real issue with the self-knowledge of the conscious will is not that it is an inferential system that can be wrong, but that the experience of conscious will seems like *an experience of a causal mechanism*, and this is impossible. “The illusion of conscious will is the belief that we are intrinsically informed of how our minds cause our action by the fact that we have an experience of the causation that occurs in our minds (2004, 682).” If we take this to be the real illusion that Wegner is attempting to spotlight, it presents some real difficulties for the bulk of his work. Not only does this directly contradict other statements about the illusion (cf...), the evidence he presents, which was analyzed in the previous section, is only ever directed at showing how the experience of conscious will has inferential mechanisms; he offers no evidence to prove we experience the conscious will as anything like a causal mechanism. If Wegner wanted to press this point and argue that the moral of automatism is that people, such as Godfrey, are only shocked by the evidence of their involvement because they believe they know how their own mind works, that is, they have experiences that tell them when they willed something or not, we would point out that there is a very important difference between an experience and a belief. In the previous quotation Wegner says that the illusion is “the *belief* that we are intrinsically informed... [italics

mine].” A false belief about mental mechanisms is different from an illusory experience.

Again Nahmias puts it well:

Perhaps, however, Wegner is better interpreted as suggesting not so much that our *experiences* of will and agency are mistaken, but that our *beliefs* about them are mistaken. That is we may be radically *mistaken* about how the relevant causal processes work...It should be no surprise that our folk psychological theories are not entirely correct, just as our folk physics and folk biology are not entirely correct. But these theories derive from more than our experiences, and our experiences are consistent with different (and more accurate) theories (2005, 778).

So in short if Wegner wants to avoid the criticism that an inferential system does not an illusion make, he needs to demonstrate not that people have incorrect *beliefs* about causal processes, but that the experience itself is about the causal process, and this is something he does not show. The combination of these two points by put so clearly by Nahmias, puts Wegner in doubtful territory. At best, and even this is debatable, Wegner has shown we have a fallible inferential system that produces an experience, but he has not shown anything about our conscious thoughts not causing our actions or even that our experience of conscious will is an experience of a causal mechanism.

Peter Carruthers: A defense

Peter Carruthers presents a defense of Wegner’s illusion argument by supporting the notion that access to the will is always interpretive. He claims,

Our awareness of our own will results from turning our mind-reading capacities upon ourselves, and coming up with the best interpretation of the information that is available to it—where this information doesn’t include our acts of deciding themselves, but only the causes and effects of those events (2007, 199).

The problem with having only interpretive access to our will is that interpretive access is by definition not immediate and direct access, and conscious experience, on Carruthers account, is always direct and immediate. If we only have interpretive access then, there is no conscious experience of the acts of the will, that is, there is no such thing as conscious deciding, therefore any experience that is taken to be a conscious experience of the will must be an illusion.

I shall take for granted that if it can be shown that the only form of access that we have to our own intentions and decisions to act is interpretative—in this respect like the access that we have to the intentions and decisions of other people—then there is

no such thing as conscious willing or conscious deciding. Conscious will would have been shown to be illusory (2007, 210).

In some ways the best response to Carruthers is simply to acknowledge that he might be right, but at the same time deny that he has actually accomplished anything by his defense. Assuming that Carruthers' account of decision making is correct and that the system of practical reasoning is not a conscious system, not in the right way anyways, we still might say that Carruthers has not shown how conscious intentions cannot cause actions, merely that we cannot have a direct experience of the process. This still leaves us with the ability to have experiences that accurately inform us as to what we are doing, to correctly own the authorship or control of our actions, or have any other experiences that are veridical aspects of our sense of agency. To prove that we can have no direct experience of an act of willing just means that *if there was such an experience* it is an illusion, but Carruthers does nothing to show that we do have this experience, as opposed to a *belief* about our experiences and psychological mechanisms.

Perhaps there is another way of putting this. Let us say that the process that produces an action involves multiple influences, only some of which are conscious. For example, walking through a room I see a book I want, and have the thought, 'Oh, I should pick that up so I don't forget it.' Then I pick it up. Undeniably this action involves a lot of non-conscious cognition, and even the fact that I did the action that my thought was about does not in itself show that the conscious thought was the one and only thing that brought about my action. But it seems to me that Carruthers' argument is that having this kind of thought is not the right kind of conscious experience to produce action on its own and that all other cognitive processes are unavailable to consciousness in the right way, but this still does not show that an illusion has occurred. It might be the case that on the basis of our conscious thought and accompanying action we form the *belief* that our thought caused the action, but this is a (possibly) false belief and not an illusory experience. And our conscious thought cannot be the illusion because it is not about anything other than its semantic content, which does not represent anything about the mechanisms of my action production. Carruthers' argument first needs to show that there is an experience that self-presents as a direct experience of mental causation, before he can show that this experience must be illusory. So even if we give up something like an experience of the

conscious will as illusory, as Carruthers wants us to, this does not mean that we have given up either the involvement of conscious thoughts in our behaviour or that *all* the experiences that are contributors to our sense of agency are not veridical.

A final comment about Carruthers' argument is suggested in work done by Elizabeth Pacherie and it nicely extends the point just made. Pacherie argues that our sense of agency does not seem to need the kind of access to mental states, such as intentions, that Carruthers is denying we have. It is not problematic for Pacherie that Carruthers' claims we are unable to be immediately aware of the elements in our practical reasoning system; our sense of agency does not seem to require that kind of access. She states,

...As several commentators (Marcel, 2003; Pacherie, 1996, 2001; Spence, 2001) have pointed out, prior intentions or awareness thereof do not seem to be necessary for the sense of agency. On many occasions, we cannot remember what our prior intentions were and yet do not disown the actions. Furthermore, many of our actions, impulsive, routine or automatic, are not preceded by conscious intentions and yet we own them (2007, 15).

So then if Carruthers' position is that we cannot have conscious experiences of willing because we cannot have direct access to those processes, we might respond by suggesting that we may not have the kind of experience he denies existing, but even despite our lack of access to practical reasoning we are still able to have the kind of experiences necessary to accurately inform our sense of agency. Carruthers in some ways fits well with Wegner then because his argument appears to be targeting a poorly defined, and perhaps non-existent target, which could be improved upon by a more refined and informed concept of conscious will.

Benjamin Libet

Wegner uses research done by Benjamin Libet to support his claim that there are physiologically distinct processes that produce our action and our experience of conscious will. Libet's literature in this area is important and Wegner is certainly not alone in his use of Libet, who has done pioneering work on neurological and physiological approaches to the mind-body problem. As a sort of interesting side note, in one of his well-known papers from 1992 Libet explains his goal as a researcher in this area, which also succinctly expresses the current thrust of consciousness research. "But experimental findings about the nature of the mind-brain relationship can at least provide

constraints to which any metaphysical theories should be asked to conform (1992, 255).” This is a good idea and is certainly expresses what Wegner is attempting to do as well. Libet is important to examine here for at least three reasons then: his approach of being informed by empirical results is necessary for proper theoretical analysis of consciousness, his publications in this area have been very influential in general, and some of his results are specifically a part of Wegner’s argument. But even using Libet does not help Wegner with his conceptual problems. The conclusions drawn from Libet’s results depend on a model of conscious control that simply does not constitute an appropriate picture of consciousness and behavior. The point that Wegner uses Libet as support for is simply this: Wegner argues that Libet’s results show how the conscious experience of willing an action is a latecomer on the cognitive scene that only appears after non-conscious mechanisms have set action in motion, and therefore our conscious experience of willing cannot have caused the action; being non-causally linked to the action, the conscious experience must be physiologically distinct. However, in order for Wegner to use Libet’s evidence he dilutes the concept of conscious will yet again to include experiences that are even more phenomenologically removed from the hodgepodge already associated with the concept, and consequently he loses the argumentative weight of these results towards the appeal of his argument.

Brain before mind

In the paper Wegner references in his argument for illusion, Libet describes how voluntary actions are preceded by neural activity called a “readiness potential” (RP), which also turns out to precede “the conscious intention to act (Libet 1985).” In many ways, the experimental construct Libet derives these results from is straightforward but it is this simplicity that hides its malformation, which is an unacknowledged and inappropriate assumption about the nature of conscious control.

In the experiment a subject is instructed to make a voluntary action at any point in time while paying close attention to the onset of the desire to make that movement. The subject reports the time at which the desire to move was felt, being corrected for by control testing, and the experimenter measures the “slow negative shift in the electrical potential” from the scalp. Readiness potential is the name given to the electrical activity

occurring widely in the brain that peaks less than a second before electromyography readings indicate muscle movement. The coordinated testing and naming of this potential was done prior to Libet by Kornhuber & Deecke (1965) and Libet simply took their research one step further by asking for a first person report of conscious experience in conjunction with measuring RP. In short, the experimental paradigm is one where subjects have agreed to “comply with a variety of instructions”, where these instructions include the expectation that the subject will “perform the prescribed motor act” and “pay close introspective attention to the instant of the onset of the urge, desire or decision to perform”. The subjects report that the inclination appeared “spontaneously”, that they “were consciously aware of their urge or decision to act before each act”, and that they “felt no external or psychological pressures that affected the time when they decided to act (530).” The conclusion Libet makes is identical to what Wegner claims, and so Wegner uses his final point for word:

Clearly free will or free choice of whether *to act now* could not be the initiating agent, contrary to one widely held view. The is of course also contrary to each individual’s own introspective feeling that he/she consciously initiates such voluntary acts; this provides an important empirical example of the possibility that the subjective experience of a mental causality need not necessarily reflect the actual causative relationship between mental and brain events (Libet 1992, 269).

There are three ways in which we might think that while this research says something about the relationship between conscious experience and action, it still does not sustain Wegner’s argument about the illusion of the conscious will, and I will address each of these below.

The first problem for Wegner is that the experience subjects are forced to report on in this paradigm does not have the same, or even that similar, a phenomenology to the experiences he has already associated with the conscious will. Consider his own analysis of these results:

These findings suggest that the brain starts doing something first (we don’t know just what that is). Then the person becomes conscious of wanting to do the action. This would be where the conscious will kicks in, at least, in the sense that the person first becomes conscious of trying to act (2000, 53).

We can see here that the sense of conscious will involved in this experience is that of “trying to act.” This is nothing like the other experiences, such as the experience of

intentionality or mental causation that Wegner refers to as being indicative of the conscious will. So while Libet may prove something about how we experience “trying to act” after the occurrence of RP this does nothing for Wegner’s argument. It is not at all obvious that when a subject reports on feeling the urge to act, we should understand this experience as having anything to do with what Wegner describes in other places as the experience of conscious will. The experience of conscious will he is trying to dismantle as illusory is, on his own account, a different experience from the one discussed by Libet.

The second problem for Wegner follows closely on the heels of this one and is another problem of the failure to differentiate experiences phenomenologically. Libet requires his subjects to report on an “urge, desire, or decision to act” before each action, but these are phenomenologically very different experiences with distinct phenomenal character. An urge, and to some degree a desire, has the quality of an internal pressure that motivates us for unknown or unclear reasons. We usually describe an urge as something we cannot quite explain or give rational reasons for. A decision is the exact opposite experience. It carries the quality of being an intentional action with specific reasons. By structuring the paradigm this way with a range of experiences that trigger a participant response, Wegner makes an assumption about the phenomenology a subject experiences, that it fits in this range and therefore we do not know exactly what it is the subject does experience or even if it properly falls into this experiential range. Even more importantly, what this ambiguity means is that it is at least possible that the experience subjects are reporting on, is *not* a conscious decision to “act now”, which is a reasonable gloss on the experience of ‘willing’ according to this paradigm, but the experience could be something else like the experience of control, where they feel their prior intention to act is being fulfilled. If this is the case, something that Libet’s experiment leaves untouched, then Libet has not shown how conscious willing is a mechanistic slowpoke but that we have experiences related to our agency that occur at different temporal points in action production and execution. So again, while Libet’s evidence obviously says something about the relation between a certain experience and an action, it is not demonstrated in this experiment that the experience it is assumed is being reported on is the one most appropriately understood as being the conscious cause of action.

Gallagher and appropriate timescales

The third problem is that there is an unacknowledged assumption at the heart of Libet's experiment and Wegner's interpretation of it, and it has to do with their model of conscious control. The experiment goes looking for a moment, a tiny slice of time that contains an occurrence of the conscious will. But why should we think that the slice of time approximately 200msec before an action is the point at which we should go looking? If we consider the paradigm again, it does not seem to be obvious to me that this tiny time frame is where we should necessarily look and Libet (or Wegner) gives us no argument to convince us. By not showing how this time slice is the crucial one to examine, we have no reason to think that the subjective reports from this time frame are the ones that will decide the issue of our experiences of conscious control. In fact, we might think that there are other points in time more appropriate to examine than 200msec before an action.

The reason why the temporal relation between RP and subjective report is not conclusive for Wegner is that the point in time where subjects are giving a report is not the time when the cognitive activity to produce the action began. Let us look at the experiment more closely. The experiment begins, as all laboratory experiments do, with the subject *agreeing* to participate and *choosing* to do some task according to direction. We have already seen that the subjects here agreed to "comply with a series of instructions" and "perform the prescribed motor act." If we were to judge when the intention to move was formed, it seems likely that our answer would be that it began when a subject *agreed* to participate and *chose* to follow the direction to make a move at an appropriate time, which happened at the start of the experiment and not slightly before the action occurred. This is what Searle (1983) called a prior intention. The fact that following the point in time when the intention was formed and before the action there was a brain event, the RP, and *then* after that a conscious experience of something like an 'urge' to move should not be surprising to us, because this seems to be a reasonable way of understanding the relation between consciousness and behavior, that is, that we consciously form intentions at some point in time prior to our actions and then allow the mediating processes to take place. For example, if we had the ability to measure the RP and collect the report of the 'urge to act' of a baseball player who makes a spectacular line-drive catch, what would

we expect to find? Probably we would find results similar to Libet's but we would not think that the subjective report of the player was a report of the conscious decision to move. He does not need to make that decision in the 200msec before the catch because he already formed the intention to move when he ran out on the field and this brings online the necessary mechanisms to satisfy his intention.

It also needs to be pointed out that we do not know what it is, exactly, that RP is measuring. Yes it occurs a certain amount of time prior to muscle stimulation in a voluntary action, but it also occurs in the same way before unintentional and involuntary actions. This information, that Wegner also notes even though he refers to RP as "the spark of will", changes the picture we have of Libet's paradigm again (Wegner 2002, 55). Knowing the ambiguous nature of what RP represents, we now might see the experiment as merely a scenario in which someone having consciously decided to participate in an experiment has an increase in brain activity (RP) prior to having an experience of something like the urge to move. Now we can also see that this picture does not convincingly show that the brain caused an action prior to our conscious experience of causing the action. The only way in which this picture is problematic for consciousness as a causal force is if you require conscious control, and our experience of it, to occur in the tiny slice of time immediately before an action, and I would suggest that this is actually counter-intuitive and does not reflect a properly nuanced model of the relationship between conscious intentionality and non-conscious mechanisms.

Because Wegner takes Libet's results to show that consciousness is cognitively slow compared the activity that 'actually' produces behavior, it seems to safe to say that he does not appreciate how conscious intentions can set actions in motion on longer timelines via non-conscious mechanisms. But this is exactly how our experience of conscious control represents our agency to us. In Libet's experiment, the subject is supposed to report on their conscious activity precisely at a time when we never typically attend to the processes producing our actions. There is no *prima facie* element of the experience of conscious will, or agency, that requires us to assume that our experiences occurring at 200msec prior to our actions are the experiences that constitute our sense of agency. Shaun Gallagher raises many similar points to these and he puts the problem with Libet's and Wegner's conclusions on this matter quite well when he says,

That most of this [bodily] control happens non-consciously is for the best. If, as in the case of deafferentation (which involves loss of proprioceptive feedback), we had to control our movements consciously or reflectively (Gallagher & Cole 1995), or if we were required to consciously represent our movements in a Cartesian mental space before we effected them in worldly space, we would have to exert great cognitive effort and slow things down to a significant degree. Libet's results, then, are no great surprise unless we think that we control our bodily movements in a conscious and primarily reflective way. The Libetarian experiments are precisely about the control of bodily movement, and even in this regard they are limited insofar as they effect an atypical involution of the question of motor control. In the experimental situation we are asked to pay attention to all the processes that we do not normally attend to...(Gallagher 2006, 116).

Gallagher's overall criticism of Libet is precisely the same one as I advocate for Wegner's. The arguments involved in both sets of research require a model of conscious control, and therefore conscious experience, that does not reflect how consciousness control *actually is experienced*. It makes no sense to think that the possible conscious control of an action in Libet's experiments occurs in 200msec before the action unless you do not pay attention to the way in which we consciously experience our sense of agency, which does not contain a phenomenal quality of being immediately prior to action. Failing to supply an argument for using this timeline that shows that we do experience our agency in this temporal space means that Wegner fails to find any support in Libet's neurological research.

Section 3 - The phenomenology of Agency

Thus far I have been critiquing Wegner's account of the illusion of the conscious will on the grounds that he does not target an appropriately defined concept of the conscious will. His model of how we consciously experience agency seems to be more of a straw man than a compelling picture of our experiences. Certainly, as Wegner is attempting to do, the contributions of science can always serve to correct our intuitive notions of agency and consciousness, but correctives that do not clearly identify the concept or folk-psychological item they are fixing do not correct anything. However, this is not to say that it is not possible for us to have illusions about our agency, and this is not to say that Wegner does not present information that deals with some of the ways in which we have illusions, the point made thus far is just that Wegner's case does not satisfy and that *his* dramatic conclusion about the illusion of our agency is overstated. Having criticized Wegner's argument, the goal now will be to redeem some of our experiences as veridical, and resist the idea that we cannot really know anything internally about how it is we come to do things. I do agree with Wegner that there are ways in which some of our self-knowledge about action causation can be wrong and that we fail to notice our fallibility, but these confusions are of a different experiential mode than the illusion Wegner announces. So even with the inclusion of possible confusion about our own agency, there is good reason to think that (1) we can know things introspectively about our agency, and that (2) our sense of agency is rooted in our action production mechanisms.

It may be helpful right at the outset of this section to explain exactly what it is that is *not* going to be proven, or even discussed really. No attempt will be made to demonstrate that we can experience mental causation, or that conscious causation is provable, or even that we can introspectively know certain potentially crucial things about our agency, such as the moment of willing or the moment of action initiation. The reason for this is that Wegner's argument is about the *experience* of conscious will, and even though Wegner's conscious will is not a very coherent or consistent concept, it clearly has to do with the experiences related to human causation or the experience of human agency. Wegner's point, ultimately, is that our agentic experiences can tell us nothing about our agency with any certainty, and it is to this that I will respond.

There is more to our sense of agency than judgments, and while Wegner assumes that agentive self-awareness is entirely doxastic in origin or outcome, this is primarily because he does not consider the phenomenology of agency such that he can appreciate the difference between agentive judgment and agentive experience. So even though Wegner employs dubious rhetoric like “We think we have selves,” I will not be trying to prove the existence of a self, but merely presenting some research that suggests our agentive experiences are not all of the type and nature that Wegner takes them to be. The question that would constitute a proper response to Wegner is therefore, what would it take for our agentive experiences to be veridical? What would they have to indicate about the world in order for us to think that they are not illusory? These questions are valuable not only because they might allow us to resist Wegner’s conclusion but also because they might lead us towards better empirical investigations that would aid in our theoretical notions of consciousness, agency, and selfhood. Ideally, then, we would like to move past Wegner and offer more sophisticated concepts and models of our sense of agency, towards which we might direct new empirical efforts.

This final section has four parts all of which attempt to clarify some of the issues surrounding our sense of agency as it is informed by our experiences of agency. The goal is to show that despite our ability to misjudge our participation in an action, taking ourselves to be agents is a justifiable belief that is supported by our experiences of agency that are in turn grounded in the mechanisms of action production. The first part (3.1) sets the stage for the others by presenting both a brief defense of the usefulness of phenomenology for these questions and a clarification of what we can know about our agency phenomenologically (Horgan 2007; Zahavi 2007; Dennet 1991). The second part (3.2) argues that there is a clear phenomenological distinction between agentive experiences and agentive judgments and that while we can be wrong about our judgments and even be mistaken in not realizing we are making judgments, this says nothing about the veridicality of our experiences (Gallagher In Press, Pacherie & Bayne 2007). Section 3.3 examines evidence (Howhy 2007; Pacherie 2007; Gallagher 2000, In Press) that suggests our agentive experiences are grounded in the mechanisms that produce action and as such we have a positive reason for thinking that our experiences of agency *are* actually about, and a part of, our agency. The final part (3.4) suggests that Wegner’s

dramatic conclusion also suffers from a misinformed picture of agent-control, and that adopting the picture of agent-control offered by Philip Pettit (2007) potentially solves some of the problems of apparent illusion, does justice to the veridicality of our agentic experiences, and makes sense of our intuitions about human agency.

3.1 – The importance of phenomenology for consciousness studies

The principle question of this section is very simply, how can phenomenology help us know about ourselves? By this, I am not assuming that phenomenology is a substitute for third person analysis by any means, but without introspecting on the nature of our experience it is impossible, or at least very difficult, to do any sort of research on the source of an experience or its role in cognition. For example, it is apparent to many people upon even the most casual reflection that there is a substantive difference in character between the experiences we have of being active in a movement and being passive in a movement, what is sometimes referred to the sense of agency (SA) vs. the sense of ownership (SO) (Gallagher 2000, In Press). The difference is the difference of feeling like an agent and feeling like a target of an external force. We ‘own’ both movements in that we acknowledge that they both occurred in relation to our body, but there is a difference in this relation, which is precisely the thing that phenomenological reflection distinguishes quite well. Phenomenology helps us divide nearby experiences along conceptual lines that enables us to do further empirical exploration more accurately. With this simple example in mind, there are then three claims about phenomenology and agency that I present in this subsection: (1) the phenomenal character of our agentic experience matters for arguments about the sources of experience, (2) there are limits to what we can know about consciousness via phenomenology, (3) appreciation of the distinctiveness of our phenomenology requires sophisticated conceptual resources. The upshot of these three claims is again that empirical investigation into consciousness, such as searching for the neural correlates of conscious experience or rational decision making, requires a richly-informed conceptual model of the relationship between experience and agency and this is best done by incorporating available phenomenological insights. Phenomenology is generally not

done well, which is precisely why Wegner employs such an awkward concept of conscious will, and is also why people in general

Phenomenology matters

I have frequently referred to Wegner's phenomenological shortcomings, but the fact that he never even mentions phenomenology despite how relevant it appears to be suggests he might be one of those who fail to find it empirically valid. It needs to be said that the idea that phenomenology matters for empirical studies of consciousness is a matter of ongoing contention (See Zahavi 2007; Dennett 1991; Metzinger 2003). Daniel Dennett, for example, argues that the true subject of consciousness studies are the first person reports that experimental subjects offer, not the entities the reports are supposedly about. He suggests that first person accounts of phenomenal experience amount to belief about phenomenology but do not demonstrate that this phenomenology exists. In a response to Dennett's 'heterophenomenology', Dan Zahavi characterizes Dennett's position on phenomenology that is scattered throughout his work in this way:

That something has subjective or experiential reality for the subject just means that the subject believes in it (Dennett, 1993b, 139). Pre-scientifically we assume that judgments are about certain phenomenal happenings, but in fact, there are no such happenings. There are the public reports we utter, there are the episodes of propositional thinking, and then there is, as far as introspection is concerned, darkness (Dennett, 1979, 95). To put it differently, there is no real seeming of yellow or inner presentation of pain over and above the thinking (judging, believing) that something seems yellow or feels painful (2007).

The problem with phenomenology is then that it less research than it is theorizing. But this attack on the usefulness of phenomenology as unscientific is misguided, says Zahavi. It assumes that phenomenology is something like introspective psychology, interested in the personal qualia of experiences that are fallible and corrigible. Zahavi says that on the contrary, "Phenomenology aims to disclose structures that are intersubjectively accessible, and its analyses are consequently open for corrections and control by any (phenomenologically tuned) subject (2007, 31)." Without going into a history of phenomenology and the aims of its chief proponents (i.e. Husserl, Heidegger, Merleau-Ponty), I think that some temporary truce can be found between these two positions, and the usefulness of phenomenology for the present discussion justified.

As mentioned previously, Wegner's work is centered on the *experience* of conscious will, and as such he invites considerations into the character or phenomenal quality of this experience. Even if it is debatable whether this experience is merely a (possibly false) belief about the presence of a certain mental entity, as Dennett might claim, Wegner's argument about what an experience can or cannot tell us demands the attempt at the carving up of experience into distinctly different, and therefore experimentally targetable, experiences. It is still an open question, subject to independent criticism and consensus, whether or not a particular experiential taxonomy is satisfying, but in principle it is broachable. I would say that the characterization of phenomenology as amounting to irresolvable disagreements about the subjective grasp of a sensory input, something like the dispute over the specific character of a colour experience (See Metzinger 2003, 591), is a hasty judgment. If it can be shown that discussions regarding phenomenal qualities can reach consensus on some of the structure, and possibly even the satisfaction conditions, of an experience, then it should be granted that the conceptual framework of an experience as informed by phenomenology is a valid structure for incorporation into other empirical studies. Besides all this, Wegner himself sets the wheels in motion for phenomenological considerations of his research by stipulating that we have an experience, the conscious, with certain experiential qualities, that of being a self-knowing mechanism. Therefore even if Dennett is right, phenomenology is important for the present discussion because the only way Wegner can make any claim about an experience being illusory is if there is some question about the contents of our experience *not* being illusory and this means taking seriously the experiential qualities of agency.

3.2 – Agentive self-knowledge: experience vs. judgments

Assuming that phenomenology does have something to offer us, many nuances in our sense of agency become apparent. John Horgan points out that we are able to extend our phenomenological insights by developing increasingly sophisticated conceptual structures. Paying close attention to the differences between the concepts of intentionality, authorship, and control allows us to distinguish between similar and related experiences of agency. Of course, as Horgan also points out, our conceptually informed phenomenology means that we can appreciate the difference between aspects of

our sense of agency that are judgmental and those that are more basic, or simply experienced. This is not an easy difference to appreciate, which is at least one reason why people experience things like automatism. We are good at looking for agency but not good at analyzing our modes of self-knowledge. There are however, at least two ways in which we might see this difference. The first of these is given by Gallagher (2000, in press) where he suggests that we can conceptually identify the difference between pre-reflective and reflective self-awareness. He uses a passage from Alvin Goldman that expresses how perception is always accompanied by this kind of pre-reflective self-awareness.

[Consider] the case of thinking about x or attending to x. In the process of thinking about x there is already an implicit awareness that one is thinking about x...When we are thinking about x, the mind is focused on x. Nevertheless, the process of thinking about x carries with it a non-reflective self-awareness (Goldman 1970, 96).

This kind of awareness might be described as being 'thin' in representational content. Not much information about the world is given by this kind of awareness, but this does not mean that there is no experience of agency present or that it necessarily includes some element of judgment. The notion of pre-reflective self-awareness is simply the idea that before the conceptual analysis of our experience begins, there is a basic awareness of self present to us as being the vehicle of the experience. This does not get us the whole way to identifying when judgment begins but it opens the door to the distinction.

Pacherie and Bayne also distinguish between the modes of agentive self-awareness and argue that the difference between doxastic and experiential awareness is apparent because of the difference between the vehicles of experience and judgment. They argue that the vehicles of agentive self-awareness are often quite primitive. The following is the argument for their claim:

Think of what it is like to push a door open. One might judge that one is the agent of this action, but this judgment is not the only way in which one's own agency is manifested to oneself; indeed, it is arguably not even the primary way in which one's own agency is manifested to oneself. Instead, one experiences oneself as the agent of this action. Such states are no more judgments than are visual experiences of the scene in front of one or proprioceptive experiences of the current position of one's limbs (2007, 466-467).

Pacherie and Bayne go on to explain that of course agentive experiences and judgments are related but not necessarily correlated, that is, we frequently make simple judgments on the basis of our experience that are no more than affirming our experience, but as Wegner shows in his discussion of automatism and ‘I Spy’ we can certainly also make judgments that are *not* bound to our experience. I should clarify at this point that I am not going to try and provide the framework for picking apart the phenomenology of our experiences of agency versus our judgments of agency, but argue that it is enough for the present discussion to show that there is an appreciable and arguable difference and that it is possible to do this picking apart. The goal is to show that not all self-awareness is judgmental, even though fallible judgments may represent a significant part of our ongoing sense of agency. It is clear that the narrative aspects of our sense of agency are informed by our judgments and that these aspects play a large part in our folk psychological notions of agency, but we can also point to experiences that basic phenomenology shows do not fit into the doxastic category thereby allowing further support for the vindication of our sense of agency.

3.3 - Sources of our experiences of agency

It is possible that Wegner would grant almost all the points argued in this section. He probably would not deny the ability of phenomenology to distinguish between the character of our experiences, or that we can even appreciate the difference between modes of agentive self-awareness. The clincher, and what Wegner is most opposed to, is to show that there is good reason to think that the most basic elements of our sense of agency are actually produced by action production mechanisms so that the reports our experience give about our action are veridical.

Possibly the best reason for thinking that our agentive experiences have a low-level source that is more basic than the cognitive act of judgment is the fact that we not only can observe dissociations of judgment and experience but that there are experiences that are not doxastically penetrable. Pacherie and Bayne show this with an example from James’ *Principles of Psychology*. It is a lengthy quotation but it provides a compelling image.

A patient with an anesthetized arm is asked to raise it. The patient's eyes are closed, and unbeknown to him his arm is strapped to the armrest and prevented from moving. Upon opening his eyes, the patient is surprised to discover that his arm has not moved, for he had experienced himself as moving his arm. He now judges that actually he has not moved his arm. Suppose the patient is asked to close his eyes again and raise his arm. Despite now knowing that his arm is strapped to the armrest and therefore judging that he is not moving his arm, he may nonetheless still experience himself as moving his arm. Simply judging that one is not moving a limb may not remove the experience of agency (Wegner 2007, 477).

Knowing that our experiences of agency are not always subject to our judgments does not yet tell us the source of this experience, but this account suggests that it is bound to mechanisms involved in the intentions and anticipated outcome of the subject.

There has been extensive research on the need for a cognitive system to identify its own actions versus the actions of another and the proposed solutions for this problem all involve low-level 'comparator' sources that are quite likely the best account of why we experience a minimal or pre-reflective sense of agency. The experiential character of a comparator source begins in the notion of the forward modeling of efference copies of motor commands (Blakemore, Wolpert & Frith 2002). Pacherie and Bayne review evidence that shows how awareness of movement is produced in the brain regions that are upstream of the primary motor cortex but downstream of the pre-motor structures, which indicates that motor plans themselves and not just afferent feedback plays a role in agentic awareness (See Haggard and Magno 1999). They also discuss studies done by Sato and Yasuda (2005) that demonstrate how the sense of self-agency is attenuated by the degree of congruency between the outcomes predicted by efference copies of motor plans and afferent feedback (See Pacherie and Bayne 2007 for more a more extensive review). The amount of research on the relationship between agentic experience and non-doxastic cognitive activity could be considered substantial enough to constitute a near consensus. This does not deny the presence and influence of narrative sources in the sense of agency, but there is good reason to think, as has been shown, that judgment does not capture all the experiential qualities of agentic experience and that the representational content of this non-judgmental source is given by the cognitive systems of intention mapping and outcome prediction.

3.4 – Agency is a property of agents not mechanisms

The final part of this section is a small one and represents a minor attempt to combine some of the claims argued thus far into a picture of an agent that offers better opportunities for empirical study. So far I have focused on the experience of agency, as this is what Wegner primarily claims is an illusion, but this still leaves the issue of agent control untouched. All I attempt to do here is to touch on a way in which we might rethink our model of agent control such that our insights into the sources of the sense of agency are more fully integrated in our story of agent control.

One of the assumptions Wegner makes about the experience of agency has to do with the model he employs of the production of action by an agent. Philip Pettit (2007) outlines this position, what he calls the act-of-will account, and suggests that it is the standard philosophical account. Wegner, in following this account, assumes that any story of how an agent causes an action must make reference to an act of the will that is an instantaneous expression of the agent-self and that we can scientifically identify the point of will-formation. Pettit however argues that the act-of-will story is precisely what leads to the kind of conclusions drawn by Wegner. By assuming that an isolatable act of the will is a necessary part of the etiology of an action, Wegner then goes looking for the experience of this act and not finding it declares the experience of agency or agent control an illusion. Pettit's response is to argue that "an action is agent-controlled, not in virtue of the elements of its particular etiology but in virtue of the nature or constitution of the agent in whom it is produced (2007, 8)." The evidence that agents have the constitution necessary to produce actions is demonstrated by their ability to respond to reason, Pettit says.

One of the nice implications of this story, is that rather than expecting action to follow a particular path of will formation and action production, we can assume a much more complicated process with multiple influences on behavior. We know that sometimes reasons influence us more than other times, but it seems intuitively obvious that they influence us at least some of the time. Our picture of the sources of the experience of will, then, fits nicely with this more complicated picture of action production. The reason for this is that the act-of-will story seems to assume something like a Cartesian inner

space that contains our mental acts and conscious experiences, like the conscious will, and having seen that our sense of agency at least partially arises from comparator sources that are not reducible to a meta-cognitive type of introspection, it seems that a complicated system of action production goes well with a complicated and sophisticated system of self-monitoring. If determining our involvement in action was just as simple as reflecting on the internal presence, or lack thereof, of an act of the will and sensory feedback regarding completion then there would seem to be no need for the deeply integrated sources of the sense of agency it appears we have. The picture of how we know about our agency that is targeted by Wegner deserves to fall, if for no other reason than it portrays consciousness as a detached Cartesian space of observation, as opposed to an integrated cognitive system that is fully embodied and functional. Why would we not think that being the sophisticated cognitive systems that we are, our conscious experiences would not also be not just part of the loop, but the loop itself.²

² This is obviously a reference to Dennett (2003), “I am the loop.”

Concluding remarks

Daniel Wegner has responded to criticism by arguing that those who fail to appreciate the truth of the illusion of conscious will do so because it is a tough pill to swallow. Our personal experience pushes against our critical science, and this, he says, cannot be undone. “This deep intuitive feeling of conscious will is something that no amount of philosophical argument or research about psychological mechanisms can dispel (2002, 325).” Having presented some serious challenges to his account, we might then wonder about the fact, admitted to here by Wegner, that we have such strong intuitions that run against his research. Rather than assuming that our intuition is fueled by the desire to see ourselves as agents, to be good agency detectors, we might think that our intuitions of conscious agency, in a broad sense, are tied to our functional success in the world. Even though we know of many instances where functional success did not translate into truth (i.e. the Ptolemaic conception of the solar system), it only seems prudent to assume that our most basic kinds of experiences are related to the systems that have evolved to help us survive rather than assuming these experiences are produced by a desire to have a certain kind of self-image. Doing empirical research into our conscious experiences then, should begin with careful considerations of our targeted concepts, doing things like proper phenomenology, expecting that while we may discover that not all of our experiences are what we take them to be, many will be redeemed.

References

- Aarts, H., R. Custers, & D. Wegner. (2005). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and Cognition*, 14: 439-458.
- Bayne, T. (2006). Phenomenology and the feeling of doing: Wegner on the conscious will. In S. Pockett, W.P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition*. Cambridge: MIT Press.
- Bayne, T., & N. Levy. (2006). The feeling of doing: Deconstructing the phenomenology of agency. In N. Sabanz & W. Prinz (Eds.), *Disorders of volition*. Cambridge: MIT Press.
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Science*, 6(6): 237-242.
- Boswell, J. (1791). *Life of Johnson*. London: Oxford University Press, 1970 (3d ed.).
- Dennett, D. (1979). On the absence of phenomenology. In D. Gustafson & B. Tapscott (Eds.), *Body, mind, and method*. Dordrecht, The Netherlands: Kluwer.
- Dennett, D. (1982). How to study human consciousness empirically, or, nothing comes to mind. *Synthese*, 53:159–180.
- Dennett, D. (1984). *Elbow Room*. Oxford: Oxford University Press.
- Dennett, D. (1987). *The intentional stance*. Cambridge: MIT Press.
- Dennett, D. C. (1991). *Consciousness explained*. Boston: Little, Brown and Company.
- Dennett, D. (1993a). Caveat emptor. *Consciousness and Cognition*, 2: 48–57.
- Dennett, D. (1993b). Living on the edge. *Inquiry*, 36: 135–159.
- Dennett, D. (2003). Who's on first? Heterophenomenology explained. *Journal of Consciousness Studies*, 10(9–10): 19–30.

- Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science, *Trends in Cognitive Science*, 4(1): 14-21.
- Gallagher, S. (2003). Phenomenology and experimental design: Toward a phenomenologically enlightened experimental science. *Journal of Consciousness Studies*, 10(9–10): 85–99.
- Gallagher, S. (2006). Where's the action? Epiphenomenalism and the problem of free will. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition*. Cambridge: MIT Press.
- Gallagher, S. (In Press). Multiple aspects in the sense of agency. *New Ideas in Psychology*.
- Goldman, A. (1970). *A theory of human action*. New York: Prentice-Hall.
- Haggard, P. & E. Magno. (1999). Localizing awareness of action with transcranial magnetic stimulation. *Experimental Brain Research*, 127: 102–107.
- Holton, R. (2009). *Willing, Wanting, Waiting*. Oxford: Clarendon Press.
- Horgan, T. (2007). Agentive phenomenal intentionality and the limits of introspection. *Psyche*, 13(1).
- Hohwy, J. (2007). The sense of self in the Phenomenology of Agency and Perception. *Psyche*, 13(1).
- Hurley, S. (2009). Bypassing Conscious Control: Unconscious Imitation, Media Violence, and Freedom of Speech. In S. Pockett, W.P.Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition*. Cambridge: MIT Press.
- Jackson, F. & P. Pettit. (1990). In Defense of Folk Psychology. *Philosophical Studies*, 59: 31-54.
- James, W. (1890). *The Principles of Psychology*. Vols. 1 and 2. New York: Dover

- publications, 1950.
- Jeannerod, M. (2006). Consciousness of Action as an Embodied Consciousness. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition*. Cambridge: MIT Press.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8: 529-566.
- Libet, B. (1992). The neural time-factor in perception, volition, and free will. *Revue de Metaphysique et de Morale*, 2:255-272.
- Maasen, S. (2003). Questioning the multidisciplinary field. In S. Maasen, W. Prinz, & G. Roth (Eds.) *Voluntary Action: Brains, minds, and sociality*. Oxford: Oxford University Press.
- Malle, B. (2006). Of Windmills and Straw Men: Folk Assumptions of Mind and Action. In S. Pockett, W.P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition*. Cambridge: MIT Press.
- Metzinger, T. (2003). *Being no one*. Cambridge: MIT Press.
- Nahmias, E. (2002). When Consciousness matters: A critical review of Daniel Wegner's The illusion of Conscious will. *Philosophical Psychology*, 15(4): 527-541.
- Nahmias, E. et al. (2004). The Phenomenology of Free Will. *Journal of Consciousness Studies*, 11(7-8): 162-79.
- Nahmias, E. (2005). Agency, Authorship, and Illusion. *Consciousness and Cognition*, 17: 771-785.
- Nahmias, E., D. Coates, & T. Kvaran. (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy*, 31: 214-242.
- Pacherie, E. (2007). The sense of control and the sense of agency. *Psyche*, 13(1).

- Pacherie, E. & T. Bayne (2007). Narrators and Comparators: The Architecture of Agentive Self-awareness. *Synthese*, 159: 475-491.
- Prinz, W. (2003). How do we know about our own actions? In S. Maasen, W. Prinz, & G. Roth (Eds.), *Voluntary Action: Brains, Minds and Sociality*. New York: Oxford University Press.
- Pettit, P. (2007). Neuroscience and Agent Control. In D. Ross, et. al. (Eds.), *Distributed Cognition and the Will: Individual Volition and Social Context*. Cambridge: MIT Press.
- Ross, P. (2006). Empirical Constraints on the Problem of Free Will. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition*. Cambridge: MIT Press.
- Sacks, O. (1970). *The man who mistook his wife for a hat*. New York: Harper & Row.
- Sato, A., & A. Yasuda. (2005). Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of ownership. *Cognition*, 94: 241–255.
- Searl, J. (1983). *Intentionality: An essay in the philosophy of mind*. New York: Cambridge University Press.
- Sellars, W. (1962). Philosophy and the Scientific Image of Man. In R. Colodny (Ed.), *Frontiers of Science and Philosophy*. Pittsburgh: University of Pittsburgh Press.
- Wegner, D. (1989). *White bears and other unwanted thoughts*. New York: Viking Penguin.
- Wegner, D. & R. M. Wenzlaff. (1996). Mental Control. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social Psychology: Handbook of basic principles*. New York: Guilford.
- Wegner, D. & T. Wheatley. (1999). Apparent Mental Causation: Sources of the

- experienced will. *American Psychologist*, 54(7): 480-492.
- Wegner, D. (2002). *The Illusion of Conscious Will*. Cambridge: MIT Press.
- Wegner, D. (2003a). The Mind's Best Trick: How we experience the conscious will. *Trends in Cognitive Science*, 7(2): 65-69.
- Wegner, D. & J. Erksine. (2003b) Voluntary involuntariness: Thought suppression and the regulation of the experience of will. *Consciousness and Cognition*, 12: 684–694.
- Wegner, D. (2004). *Precis of The Illusion of Conscious Will [and Commentaries]*. *Behavior and Brain Sciences*, 27(5): 634-692.
- Wegner, D. (2005). Who is the controller of controlled processes? In R. Hassin, J. Uleman, & J. Bargh. (Eds.), *The New Unconscious*. Oxford: Oxford University Press.
- Zahavi, D. (2007). Killing the straw man: Dennett and Phenomenology. *Phenomenal Cognitive Science*, 6: 21-43.