Against the Illusory Will Hypothesis

A Reinterpretation of the Test Results in Danial Wegner and Thalia Wheatly's *I Spy* Experiment

Robert Reimer

Universität Leipzig, Fakultät für Philosophie,
Beethovenstraße 15, 04107 Leipzig, Germany
robertreimer@gmx.de

Abstract. Since Benjamin Libet's famous experiments in 1979, the study of the will has become a focal point in the cognitive sciences. Just like Libet the scientists Daniel Wegner and Thalia Wheatly came to doubt that the will is causally efficacious. In their influential study *I Spy* from 1999, they created an experimental setup to show that agents erroneously experience their actions as caused by their thoughts. Instead, these actions are caused by unconscious neural processes, so that this 'causal experience of will' is just an illusion. Both the scientific method and the conclusion drawn from the empirical results have already been criticized by philosophers. In this paper, I will analyze the action performed in *I Spy* and criticize more fundamentally the assumption of a 'causal experience of will'. I will argue that the *I Spy* study does not show that the agent's causal experience of will is illusory, because it does not show that there is a causal experience of will. Against Wegner and Wheatley's assumption, I will show that it is unlikely that the participants in *I Spy* experienced their conscious thoughts as causally efficacious for an action, that they did not perform at all. It is more likely, that they experienced their own bodily movement as causally efficacious for a cooperative action, that they did not perform solely by themselves.

Keywords: Philosophy of cognition, consciousness of will, causation in action

Introduction

In their paper *Apparent Mental Causation* (1999), Daniel Wegner & Thalia Wheatly write: "Conscious will is a pervasive human experience." (Ibid.: 480) However, after all, it might be an illusion. Specifically, it might be an illusion, that an agent's conscious thoughts to perform an action cause that very action.

The neuroscientists Wegner & Wheatly are proponents of the illusory will hypothesis. According to it, agents only have the impression that there is a causal relation between their thought of performing an action and the very performance of that action. They call the experience of this causal relation 'causal experience of will'. Instead of being caused by her own thoughts, the agent's action is caused by unconscious neural

processes. The illusionary causal experience of will arises, because these unconscious neural processes that actually cause the action, additionally produce the conscious thoughts about it, which remain accompanying and impotent epiphenomena.

To refute the assumption, that there is a real causal relation between the agent's thoughts of performing an action and the very performance of that action, Wegner & Wheatly tried to create an experimental setup, in which participants first develop a thought to move an object and second perceive a movement adequate to this thought (the supposed action), without actually moving the object. If it further turned out that the participants in the experiment still have the impression that they moved the object, Wegner & Wheatly believed to show that we (in general) just have the "[…] feeling we willfully cause what we do." (1999: 490) This experiment was called 'I Spy'.

Wegner & Wheatly are not the first scientists setting up an experiment to test the causal efficiency of the conscious will. The pioneer scientist in the field of consciousness, Benjamin Libet, also assumed that the conscious will must be something that causes the action to happen and that can be felt by the agent (1999: 49). In a series of experiments in 1979, Libet advised the participants to measure the time when they felt the will to act in advance of acting itself. After measuring their neural processes, Libet discovered that their will to act was preceded by unconscious nerve cell activities in the motor cortex. Based on that Libet concluded, that the will is either an epiphenomenon (and therefor causally impotent) or caused by these previous activities (and therefor itself causally determined) (54). Wegner & Wheatly's hypothesis about the nature and function of the will is highly influenced by Libet's experiments.

On the other hand, Wegner & Wheatly's *I Spy* study and their theory about the nature of the will influenced the work of many other scientists in the field of psychology and neurobiology such as Lau (2006), Haggard (2008), Mogi (2014). Haggard, for example, not just shared the assumption of a distinction between what he called 'the experience of intention' and 'the experience of agency' (2008: 941-942). He also defended the illusory will hypothesis. Just like Wegner & Wheatly, Haggard tried to create situations in which agents have thoughts about an action without, actually, performing the action.

Wegner himself further developed his illusory will hypothesis in his book *The Illusion of Conscious Will* (2002) and further defended it in other, similarly structured studies, such as the *helping hands experiment* (see Wegner et al. 2004).

In this paper, I will take a closer look at the 'causal experience of will' that is presupposed by Wegner & Wheatly's illusory will hypothesis. I agree with Wegner & Wheatly that there is no causal relation between the thoughts (of willing to act) and that very action. However, I do not think that there is an illusion of such a causal relation, either. In acting, agents do not experience any causal relation between a previously experienced thought and a subsequent action. Acting and willing to act are not two separable and causally dependent events. They are synchronous and inseparable. Given that this is true, Wegner & Wheatly's illusory will hypothesis is also refuted.

Many philosophers such as Wittgenstein[1] have already argued against the assumption of a separate experience of will and in favor of a unity between thinking and acting. I will not participate in this general discussion within the philosophy of action. However, my aim is to support the assumption of the unity between thinking and acting indirectly by showing that the test results of *I Spy*, differently interpreted, do not support the assumption that there is a causal experience of will.

In section (1), I will briefly explain what it means to experience a causal relation in general and what it could mean to experience a causal relation between one's own thoughts of willing to perform a certain action and that very action. In section (2), I will introduce the experiment, which is supposed to show that agents can have thoughts and experience these thoughts as (more or less) causally efficacious for the action without acting. In section (3), I will show that the empirical results do not support that interpretation. Instead, I will present my own alternative interpretation, according to which the participants in *I spy* rather moved their fingers and experienced these bodily movements as (more or less) causally efficacious for the execution of the overall action.

1 The Experience of Causal Relations

Consider the following situation: You are playing billiard. You hit the white ball and observe it hitting the black ball. Then you observe the black ball starting to roll, too. You assume that these events are causally dependent on each other, such that the white ball by striking the black ball causes the black ball to roll, too:

$$\text{the white ball (by hitting the black ball)} \xrightarrow[causes]{} \text{the black ball (to roll)}^2$$

Even if you are skeptical at the beginning whether there is a real causal relation between these events, your skepticism will vanish after repeating the trial.

David Hume argued that the idea of a causal relation (or a causal connection) is based on the perception of the temporal succession of two events that can be repeated. If you notice that events of type A (e.g. the movement of the white ball) are regularly followed by events of type B (e.g. the movement of the black ball), your mind concludes that there must be a causal relation between instances of type A and instances of type B. But according to Hume you just suffered from a 'causal illusion' in postulating that relation. There was no causal relation between these events, just a succession of them.

---

[1] Consider this passage in the *Philosophical Investigations*: "When I raise my arm 'voluntarily', I do not use any instrument to bring the movement about […] 'Willing if it is not a sort of wishing, must be the action itself […]'" (1999, 160e).

[2] Each causal relation consists of two objects (agens and patiens) being involved in two separate events such that the event of the patiens (effect) is causally dependent on the event of the agens (cause). I will frequently use schemas like this to illustrate the structure of certain causal relations. These schemas should be read in the following way:

$$\text{agens (cause-event)} \xrightarrow[causes]{} \text{patiens (effect-event)}$$

For Hume, not only events in the physical world give rise to such a ‘causal illusion’, but also the ‘acts of the spirit’. He wrote:

“Some have asserted, that we feel an energy, or power, in our own mind; […] But to convince us how fallacious this reasoning is, we need only consider, that the will being here consider’d as a cause, has no more a discoverable connexion with its effects, than any material cause has with its proper effect. So far from perceiving the connexion betwixt an act of volition, and a motion of the body; […] the actions of the mind are, in this respect, the same with those of matter. We perceive only their constant conjunction; nor can we ever reason beyond it.” (1978: 633)

Wegner & Wheatly, in referring to Hume, did not want to support his general skepticism of causal relations. In fact, their theory rests on the assumption that actions are caused by unconscious neural processes. However, they adopted the skepticism of mental causation and two central ideas:

First, they noted that causal relations in general cannot be perceived directly. Causality is not a ‘magic bond’ between events that can be made visible under the microscope. Instead, perceivers must infer a causal relation based on the experience of the repeated succession of two events of a certain type. That explains why causal theories are prone to illusions: There is no proof whether events of two types really stood in a causal relation to each other or whether one was just accidentally followed by the other.

Second, and more importantly, Wegner & Wheatly also shared the assumption with Hume that the perception of events in the physical world basically resembles the experience of one’s own agency. They wrote:

“The person experiencing will […] is in the same position as someone perceiving causation as one billiard ball strikes another. Causation is inferred from the conjunction of ball movements, and will is inferred from the conjunction of events that lead to action.” (1999: 480)

Note, that according to this analogy, the object of perception in case of agency must be identical to the subject of perception. It is the agent who is supposed to experience a causal relation ‘within’ herself. First, she experiences some of her own thoughts when they ‘occur in their consciousness’, as Wegner & Wheatly put it (1999: 484). These thoughts constitute the supposed cause of the supposed causal relation. Then, she experiences her own action either through observation of her limbs or through proprioception. This action constitutes the supposed effect of the supposed causal relation. Based on the experience of the thought processes and the action, the agent gains the impression that her thoughts have caused the action (482). Or to put it in the words of Wegner: *“[P]eople experience conscious will when they interpret their own thought as the cause of their action*.” (Wegner 2004: 654)

However, there is no proof of a causal relation between thought and action, even if both occurred and match each other. The action could also have been caused by something else – namely an unconscious neural process. If that is true, the power of will is a causal illusion. (See Fig. 1)

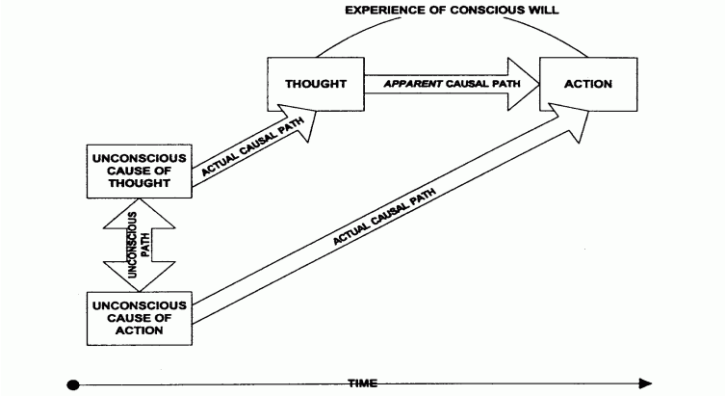

Fig. 1. Wegner & Wheatly used this schema in (1999) to illustrate the real and the apparent causal path from the unconscious neural process to the action according to their hypothesis (483).

2 The *I Spy* Experiment

Let me now briefly explain the experimental setup of *I Spy* as well as Wegner & Wheatly's interpretation of the test results.

The experiment included a series of trials. In each trial of the experiment a participant was paired with an assistant working for the experiment's administrator. Both placed their fingers on an Ouija-board-like mouse and moved it together in a circular manner. The movements of the mouse were projected to a monitor, where the mouse cursor moved over several items on the screen (a swan, an umbrella, etc.). The participants were instructed to stop their movements freely at some point. During each trial, they listened to music through headphones (See Fig. 2). Frequently, words were spoken through the headphones of the participant that designated an item on the screen (e.g. "Swan!"). These words were supposed to prime a thought about the respective item 'in the participant's consciousness'.

Fig. 2. The picture shows a participant and an assistant moving the Ouija-board-like mouse in the *I Spy* experiment (Wegner & Wheatly 1999: 489).

The participants, erroneously, assumed that the assistants were also participants, just like them. In fact, they received secret instructions from the administrator either to stop the mouse by themselves on a specific item or to let the participants stop the mouse freely wherever they want. Those trials, in which the assistant was instructed to let their participant stop freely were called 'free stops'. Those trials, in which the assistant was instructed to force a stop on an item, were called 'forced stops'. In case of the forced stops, the participants heard the word for the respective item through their headphones 30, 5 or 1 second before, or 1 second after the mouse was forced to stop on the item.

After each trial, the administrator asked the participant to rate the 'level of intentionality' that they felt when the mouse stopped, both in case of the free stops and in case of the forced stops: "[T]hey each would rate how much they had intended to make each stop, independent from their partner's intention." (1999: 478) At the end, the participants rated the free stops in average 56% intentional and the forced stops in average 52% intentional. However, there was a fluctuation in the perception of intentionality depending on the time difference. Given that the primed word was spoken 1 second before the forced stop, participants rated the stop up to 65% intentional. So, it seems that even in case of the forced stops, when the assistant was instructed to move and to stop, the participants felt (at least to some extent) that they intended to make a stop.

Based on these results, Wegner & Wheatly hypothesized that the occurring thought of a certain item on the screen together with the subsequent perception of the mouse stopping on that very item, made the participant believe that their own thoughts caused their hand to move the mouse towards the primed item, even though they did not (490).

3 On the Nature of the Causal Illusion in *I Spy*

Wegner & Wheatly's study as well as their illusory will hypothesis have been criticized by many philosophers.

In his paper *Willusionism, epiphenomenalism and the feeling of conscious will* Sven Walter (2014), identified various problems of the *I Spy* study, including the low significance of the test results and the ambiguous responses of the participants that do not necessitate the conclusion that Wegner and Wheatly drew. I will return to the las point in a minute. Schlosser (2012), furthermore, remarked that there is plenty of empirical evidence supporting the assumption that the sub-personal correlates of an agent's intention are, actually, causally efficacious for her action. Carruthers (2010), arguing from the opposite direction, doubted the existence of a universal causal experience of will. He discussed a variety of empirical evidence showing that many agents (especially children and patients suffering from autism) can experience agency without experiencing these actions as caused by their intentions.

I am, clearly, sympathetic to Carruther's position. However, I want to go further and assume that even the participants in the I Spy study did not experience their intentions (or thoughts) as (more or less) causally efficacious for their action. To show this, I will provide an alternative and more plausible interpretation of the test results.

Note that the question of the action's level of intentionality, or the question "How much did you intend to make the stop?" that Wegner & Wheatly asked the participants can be interpreted in different ways: Given that the participants developed the thought of an item on the screen after hearing the respective word, it can be interpreted in the following way: "Did my thought of the primed item cause my hand to move the mouse to this item or not?" or in that way: "How strongly did my thought of the primed item cause me to move the mouse to this item?". It is clear, that this is the interpretation that Wegner & Wheatly had in mind when they asked their question. The schema of the causal experience according to this interpretation would be the following:

$$\text{my thoughts (by occurring in my consciousness)} \xrightarrow[caused]{} \text{the mouse (to move)}$$

However, the question of the action's level of intentionality could also have been interpreted in a different way by the participants, namely as the question whether or how strongly the thrust of their hand contributed to the movement of the mouse to the primed item. That interpretation presupposes a distinction not between two different events – the mental act of thinking to move the mouse and the action of moving the mouse – but within the action itself, namely between the participant's hand movement and the movement of the mouse. If the participants suffered from a causal illusion, that illusion would not consist – as Wegner & Wheatly assumed – in misjudging the causal impact of the thought on the movement of the mouse but in misjudging the causal impact of the hand's movement on the movement of the mouse. The schema of the causal experience according to this alternative interpretation would be the following:

$$\text{my hand (by pushing the mouse)} \xrightarrow[caused]{} \text{the mouse (to move)}$$

To illustrate the difference between both causal schemas, consider this example given by Wegner in his book *The Illusion of Conscious Will*. Wegner describes a situation in which he sat in front of a gaming machine in a toy store. While he moved the joystick, "[a] little monkey on the screen was eagerly hopping over barrels as they rolled toward him." (2002: 9). He was under the impression of playing a video game but in fact, the game just showed a pre-game demo. Wegner concluded: "I thought I was doing something that I really did not do at all." (2002: 10) and assumed that operating the gaming machine is a good real-life to proof his illusory-will hypothesis.

But it is not, and O'Connor (2005) and Walter (2014) have already criticized Wegner's interpretation of the situation. Wegner did do something, namely moving the joystick. He did not erroneously assume that his thoughts caused an action. He erroneously assumed that his hand (in moving the joystick) caused the machine to operate. He did not err about acting at all, but about the outcome of his moving the joystick (O'Connor 2005: 224) or about the causal effects of his moving the joystick (Walter 2014: 2224).

Unfortunately, we cannot easily apply the reinterpretation of the gaming machine situation to the *I Spy* experiment. Since the action that Wegner performed at the gaming machine involved a machine-interaction, as O'Conner put it (2005: 224), it is not comparable with the action that the participants performed in *I Spy*. Successfully interacting with a machine requires the functioning of many independently operating devices and

gears that are partially hidden. Since some of the devices and gears in the machine did not operate as they would operate if the game was running, Wegner's movement of the joystick did not have the expected outcome. He was prone to a causal illusion because he did not have direct control over or insight in the mechanism of the gaming machine.

The causal illusion of the participants in *I Spy*, however, cannot be explained in the same way, because the action did not involve any defective machine-interaction. The Ouija-mouse and the connected monitor operated faultlessly, and the movement of the mouse was transmitted properly to the movement of the cursor on the screen. It seems, that the participants had direct control over the movement of the cursor over the whole timespan. So, how could they suffer from a causal illusion? How could they have been mistaken about the effects or the outcome of their hand movement?

Even though the action that was performed in *I Spy* did not involve a defective causal mechanism, it was not a simple bodily movement, either. Since both the participants and the assistants were invited to move the Ouija-mouse in circles for the whole time, the participants and the assistants performed a cooperative or joint action. With regard to the cooperative character of the action in *I Spy*, Sven Walter noted that "[i]n cases of joint action, however, you always have to try to respond to cues from the other in order to coordinate your movements with the other movements." (2014: 2238, FN 15) Accordingly, it is likely that in case of the free stops, the assistants took part in moving the mouse to an item, along with the participants; and in case of the forced stops, it is likely that the participants took part in moving the mouse, along with the assistants, to the primed item. Or in other words: During all trials, the overall action (of moving the mouse to an item) was influenced by the thrust of both the assistant's and of the participant's hand, because both who were alternatingly pushing forward and responding to a push. This essentially cooperative character of the action during both kinds of trials can be supported by the fact that the 'level of intentionality' that was rated during both trials was roughly the same, namely 56% vs. 52% intentional.

Given that this is true and both the participants and the assistants contributed to the movement of the mouse by alternatingly pushing and by responding to a push in the forced cases, it is likely that the participants could not tell for sure to which extent the assistants actively intended to move the mouse towards the primed item: "Did the other person independently move towards that item or did they simply try to coordinate their movement with mine?" That applies to both the forced stops and the free stops, because the participants did not know, which trial was a free and which trial was a forced stop. Furthermore, given that the participants could not precisely estimate the assistants' causal impact on the mouse movement, they could also not tell for sure, to which extent they themselves actively contributed to the overall movement of the mouse, either.

Consequently, it is possible that, both in case of the forced and in case of the free stops, the participants either underestimated or overestimated their own contribution to the overall movement of the mouse. That the participants were not able to determine their own contribution to the overall movement exactly, can be supported by the indecisive answers that they gave when they were asked to rate the 'level of intentionality' for their action. In case of a non-cooperative action, an agent would exactly know if she

intended it or not. There would not be something in-between. Either she intended it entirely or not at all. Cooperative actions do not involve this certainty. That is why neither in case of the free stops nor in case of the forced stops, the participants rated the level of intentionality 100% intentional (fully intentional) or 0% (not intentional at all).[3]

I conclude that the test results of the *I Spy* experiment do not show that the participants experienced a (more or less) causally efficacious will, separate from their action. It is more likely that they experienced their own hand movement as (more or less) causally efficacious for the overall cooperative movement of the mouse. Furthermore, the test results do not support Wegner and Wheatley's illusory-will hypothesis. It is indeed possible that the participants suffered from some kind of causal illusion. But the causal illusion that they suffered from, most likely, did not consist in misjudging the causal impact of their own thoughts on the action (of moving the mouse to the primed item). It is more likely that the causal illusion consisted in misjudging the causal impact of their own hand's thrust on the overall cooperative movement of the mouse.

References

1. Carruthers, Glenn: *A problem for Wegner and colleagues' model of the sense of agency*. In: Phenom Cogn Sci 9 (3), 341–357 (2010).
2. Haggard, P.: *Human volition: Towards a Neuroscience of Will*. Nature Reviews. Neuroscience 9 (12), 934–946 (2008).
3. Hume, D.: *A treatise of Human Nature*. 2. edn. Clarendon Press, Oxford (1978) [1739].
4. Lau, H. C.; Rogers, Robert D.; Passingham, Richard E.: *On measuring the perceived onsets of spontaneous actions*. Journal of Neuroscience 26 (27), 7265–7271 (2006).
5. Libet, B. W.: *Do We Have Free Will?* Journal of Consciousness Studies 6 (8-9), 47-57 (1999).
6. Mogi, K.: *Free will and paranormal beliefs.* Frontiers in Psychology 5 (281) (2014).
7. O' Connor, T.: *Freedom with a Human Face*. Midwest Studies in Philosophy (29), 207-227. (2005).
8. Schlosser, Markus E. (2012): *Causally efficacious intentions and the sense of agency: In defense of real mental causation.* Journal of Theoretical and Philosophical Psychology 32 (3), 135–160.
9. Wegner, D. M.: *The Illusion of Conscious Will*. MIT Press, Cambridge MA (2002).
10. Wegner. D. M.: *Précis of the illusion of conscious will*. Behavioral and Brain Sciences (27), 649-694 (2004)
11. Wegner, D. M.; Sparrow, B.; Winerman, L.: *Vicarious agency: experiencing control over the movements of others*. J Pers Soc Psychol, 86 (6), 838-848. (2004)
12. Wegner, D. M.; Wheatley, T.: *Apparent Mental Causation: Sources of the experience of Will.* American Psychologist. 54 (7), 480–492 (1999)
13. Wittgenstein, L.: *Philosophical Investigations.* G.E.M Anscombe (transl.) 2. edn. Blackwell, Malden MA (1999) [1953].

[3] Walter, similarly, pointed out the unusual indecisiveness of the participants: "If 100 corresponds to 'I intended to make the stop,' then a stop that was experienced as intended should receive an average of 100, not of 56. Therefore, the fact that free stops received an average rating of 56 does not show that the correct rating for intended stops is 56, but that the free stops were not perceived as fully intended." (2014: 2227, FN 12)