CHAPTER 6

Actions and Intentions

Sofia Bonicalzi

1 Introduction

Results in the cognitive neuroscience of volition and action have been often dismissed as ultimately irrelevant, or too week at best, to legitimately tackle the philosophical issues of free will and intentional agency. By contrast, this chapter seeks to promote a more constructive perspective regarding how philosophy and cognitive neuroscience can jointly improve our comprehension of intentional agency.

The chapter is divided into seven sections. In Section 2, I present the causal theory of action as the best attempt to provide a reductive philosophical characterisation of intentional action, introducing some early and ongoing debates concerning the causal role of conscious mental states. In Section 3, I discuss how specific problems for the understanding of intentional agency, as inherited from the causal theory, originate from widely discussed pieces of empirical evidence on how voluntary processes unfold. In Section 4, I go through some of the counter-arguments that have been put forward in order to defend the classic view of intentional agency. To a various extent, these counter-arguments target the lack of ecological validity of widely employed experimental paradigms. In Section 5, I present counter-arguments of a different type, which are based on the underlying claim that no clear causal link between unconscious neural antecedents and actions can be established on the basis of neuroscientific data.

The preoccupation expressed by some of these criticisms is shareable. Nonetheless, I will suggest that the following argument is unwarranted: *Because* it does not straightforwardly rule out the causal theory, the neuroscientific angle is irrelevant to understanding intentional agency. In Section 6, I in fact argue in favour of a different approach concerning the relation between experimental research and philosophical analysis. In particular, I suggest that the former does not simply have the role of validating the latter, but plays a more constructive part in defining the common research target. I articulate these claims with some proposals and examples (Subsections 6.1 and 6.2). Some final remarks are presented in Section 7.

 | DOI:10.1163/9789004409965_008

2 Ups and Downs of the *Causal Theory of Action*

The problem of naturalising intentional agency, i.e., developing an account of intentional agency that matches the scientific view of the world, remains a crucial challenge for the philosophy of action. In this respect, the *causal theory of action* (Davidson 1963; Searle 1983; Setiya 2007) has taken the lion's share, and has been considered the most suitable strategy for naturalising intentional agency in a reductive fashion. Notoriously, its central tenet is that actions are intentional if they are appropriately caused by the right mental states, such as desires-beliefs pairs (Dretske 1988) or temporally extended plans (Bratman 2000), working as the reasons for those actions. Intentional actions are thus intrinsically complex compounds, constituted by bodily movements and conscious intentional states. Indeed, the presence of intentional states distinguishes intentional actions from mere bodily movements, such as unintentional reflexes or spasms. A basic requirement in the architecture of the causal theory is that intentional states are causally relevant in virtue of their specific content (Pacherie 2011), being both relatively abstract or context-independent and linked to the overt bodily movement in a motor specific way: It is the content of my conscious abstract desire to drink water that causes the movement of my arm towards that glass on the table.

The causal theory has been held as widely appealing on at least three different levels. First, by positing an event-causal link between mental states and the corresponding actions, the causal theory does not imply the adoption of what Strawson (1962) referred to as *panicky* metaphysical views, such as agent-causal accounts (i.e., the agent, as an irreducible substance, causes events (O'Connor 2009)) or non-causal accounts (i.e., free intentional actions are uncaused or at least non-deterministically caused (Ginet 2008; Sehon 2016)). Given such a reductive aim - mental states are meant to be somehow instantiated by physical states -, the causal theory has been also seen as a good heuristic device in order to formulate psychological and computationally tractable models of intentional behaviour (Meltzoff 1995; Haynes 2007). Second, by stipulating the causal origination of intentional actions from conscious mental states, the causal theory is compatible with the seemingly transparent phenomenology of experience, whereby intentional actions are perceived as distinct from involuntary and automatic movements. Indeed, the subjective feeling of a connection between mental states and external sensory feedback, through our bodily actions, is thought to decisively contribute to the sense of agency and control over the external world (Haggard 2008).

Third, the causal connection between conscious mental states and overt bodily movements, stipulated by the causal theory, displays an immediate relevance for ethical, normative, and legal practices. In particular, the intentional bodily actions that we are supposedly in control of, as derived from conscious intents, are recognised as specific targets of evaluative considerations, including responsibility attributions (Lagnado & Channon 2008). These three levels are tightly integrated into a unified picture, whereby intentional actions, as caused by conscious mental states, elicit the subjective sense of agency and are appropriately reached by normative evaluations. In this chapter, I mostly focus on the first of these three intertwined levels, concerning the nature of the link between conscious mental states and actions. Indeed, despite its immediate intuitiveness, the simple idea that mental states are causally involved in the production of the corresponding actions has proven to be deeply problematic.

The causal theory falls prey of many criticisms. In particular, according to a widely discussed objection (*disappearing agent*), by focusing exclusively on the causal role of mental states, the causal theory is intrinsically unable to provide a distinctive role for the agent *qua* agent as the key factor in action production. Within the causal theory, the agent is reduced to a passive bystander, unable to control the string of events that happens to take place within her mental arena (Velleman 1992). The key defence put forward by the causal theorist is exactly that positing a specific role for the agent is not needed in order to distinguish intentional actions from bodily movements. Paramount to the success of the causal theory is hence specifying the nature of the causal link between conscious mental states and actions. However, the attempt to provide an adequate account of such a causal connection has not gone unchallenged. The early and ongoing debate on deviant causal chains (Davidson 1973; Mele 1997; Schlosser 2010) highlights how conscious mental states can cause the agent's behaviour in an indirect way: By triggering a bodily movement that allows the agent to achieve the very same goal she would have wanted to achieve, but without also causing the corresponding intentional action. Typical scenarios feature agents who desire to achieve a self-posited goal but, due to the psychological state elicited by the mental state they entertain, accidentally execute a bodily movement that realises exactly the same goal. The problem is that, to properly contribute to the action, the relevant mental state must work as the reason, not just as the aetiological explanation, for the action (Anscombe 1957). To move forwards, the causal theorist has thus to clarify not only the nature of the single ingredients of the compound (mental state and bodily movement), but how these two separate elements enter into a proper relationship forming an instance of intentional agency.

Furthermore, qualifying the nature of such a relationship is crucial for compatibilist theories of free will and responsibility, often relying on the theoretical structure of the causal theory in order to justify agential control in the absence of agent causation (McKenna 2011). In particular, according to most actual-sequence compatibilist (Haji 1998) and semi-compatibilist (Fischer & Ravizza 1998) views, the agent is acting freely and/or is responsible for her action, if she acts intentionally, i.e., if she acts on the basis of the mental states constituting her reasons for action. Crucially, actual-sequence accounts contend that, for an action to be intentional, free will is not required at the level of the selection between alternatives: I am freely and/or responsibly grasping that glass of water to the extent that I intended to do so, independently of whether I could have done something else. Acting intentionally is thus often implicitly equated to acting voluntarily, namely acting under the control of the conscious (free) will. The limit of this equation emerges in the case of actions that are voluntary under a certain description (i.e., actions that are not externally imposed, differently from actions performed under coercion or duress), without being intentional in the sense prescribed by the causal theory (i.e., they are not appropriately caused by conscious mental states). Deviant causal chains constitute a key example, but this ambiguous status extends to more ordinary cases, such as instances of negligent behaviour, habitual actions, absent-minded behaviour, and possibly episodes of weakness of will (Arpaly & Schroeder 1999). Indeed, compatibilism has traditionally had a hard time in establishing whether the agent might be held responsible for willed, but not properly intentional, actions: It is unclear whether somebody can act intentionally when some sort of high order mental state does not exercise a form of authority over the corresponding behaviour.

An anti-causalist (Frankfurt 1977) having a deep influence on the development of the causal theory (Aguilar & Buckareff 2010), Frankfurt provides a hierarchical account of the authoritative role of conscious mental states in his work on the concept of a *person*. Intrinsic to the concept of a person is the capacity to identify (with) higher order mental states (i.e., the desire to have the desire to drink water) ideally exercising control over the lower level mental states (i.e., the desire to drink water) that are ultimately responsible for the agent's behaviour. In the sequence leading to action production, the authority of conscious mental states is thus directly rooted in the agent's psychological make-up: For an agent to qualify as a person, she must be able to identify (with), or endorse, the mental states from which her actions stem. The epitome of the agent who is not acting on the basis of her own free will is the unwilling drug addict who succumbs to her (*external*) desire to take drugs, and is unable to regulate her behaviour in accordance to her higher order mental states (i.e., the desire not to desire to take drugs) (Frankfurt 1971). On a similar vein, Fischer suggests that an agent is acting freely (*guidance control*), and thus responsibly, when she is acting in accordance to what she wants to do (Fischer 1999). In the next sections, I will label this overriding force of mental states over bodily actions, central to the causal theory, as *intentional action control.*

3 Psychological and Neuroscientific Arguments against Intentional Action Control

In the last few decades, the ability of the causal theory to support intentional agency has been challenged beyond the sphere of the philosophy of action, but once more in a way that questions the nature of the relation between conscious intentions and the corresponding bodily movements. The hallmark of the debate is the attempt to reconcile the claim that conscious intentions must play a role in causing actions with a naturalistic understanding of how cognitive processes unfold. Indeed, a number of findings in cognitive neuroscience have been read as potentially undermining intentional action control. If intentional action control is meant to be crucial for intentional agency, the conclusion would be that the same arguments ruling out intentional action control would *ipso facto* rule out intentional agency. In this chapter, I mostly side with those who contend that there is not enough empirical evidence for denying intentional action control. At the same time, I argue that the empirical investigation might have a different role to play in widening our understanding of intentional agency.

Benjamin Libet's pioneering research on the timing of conscious intentions marked a turning point in the debate about action initiation (Libet et al. 1982; Libet et al. 1983). In Libet's classic study, participants were asked to flex their right wrist or the fingers of their right hand at their own pace, executing a movement (one per trial) whenever they felt a wish or an urge to do so. The paradigm thus consisted in asking participants to spontaneously decide when to move, while the specific movement they had to perform was set from the beginning of the experiment. At the same time, participants were also required to watch a clock with a dot circling around it, and to remember the time (W) when they felt the desire or the urge to make the spontaneous movement. At the end of each trial, participants reported the time W they previously memorised. The beginning of the muscular motion associated to the onset (M) of the overt bodily movement was measured by an electromyogram (EMG), while the electrical activity during movement preparation and execution was recorded through an EEG system. The electrical readings from the scalp revealed the presence of the so-called *Readiness potential* (RP), a gradual rise in the electrical activity in the motor cortex and in the supplementary motor area of the brain in the second or so preceding the occurrence of an intentional action (Deecke 2000). For non pre-planned actions, the RP began about 550 ms (Type II) before M, and 350 ms before W. Where the action was pre-planned, the RP (Type I) was registered as beginning around 1000 ms before M. Results comparable with those obtained by Libet have been more recently obtained with the aid of different techniques, ranging from fRMI (Soon et al. 2008; Soon et al. 2013) to implanted electrodes (Fried et al. 2011). The results showed an even earlier onset of the brain activity preceding the action, beginning up to 10 s before the subjective awareness of wanting to make a movement (Soon et al. 2008).

Prima facie at least, the aforementioned findings provide a picture of intentional agency that does not match with the causal theory. The most radical theoretical conclusion one may draw from the neuroscientific data is some version of *epiphenomenalism*. This is the theoretical thesis that seemingly causally relevant conscious processes, such as intention formation or decisions, do not play a causal role in the initiation of the corresponding action (Nahmias 2014). An epiphenomenalism of this sort is notoriously advocated by psychologist Daniel Wegner. He argues that, if the time of conscious awareness is subsequent to the onset of the unconscious neural determinants of the action, conscious mental states cannot be the cause of the corresponding bodily movements. The apparent causal relevance of conscious mental states is no more than a by-product of *post-hoc* confabulatory inferences regarding action initiation. To support the thesis, Wegner discusses a wide array of experimental data, suggesting a disconnection between the conscious experience of acting and the actual performance of a bodily movement (Wegner 2002).

Empirical evidence has indeed shown that people are less able to monitor their motor performance than the phenomenology of agency suggests (Fourneret & Jeannerod 1998). In particular, after brain stimulation of motor areas, people may report having experiences of moving where no real movement occurred, or carry out actions without perceiving any sense of agency (Desmurget et al. 2009). Furthermore, the subjective sense of agency can be altered when intentional actions are modulated by external guidance (priming), typically leading people to overestimate their own self-efficacy (Linser & Goschke 2007; Moore, Wegner, & Haggard 2009). In addition, decades of research in cognitive and social psychology have highlighted that apparent rational conscious choices are often the result of automatic, non-intelligent, processes accompanied by little cognitive elaboration (Doris 2016; Nisbett & Wilson 1977). Taken together, these pieces of evidence undermine the clear-cut picture of intentional action as portrayed by the causal theory, by casting doubt on intentional action control. Far from being a tightly integrated, phenomenologically transparent, construct, intentional agency might thus turn out to be a black box.

4 Counter-arguments Based on the Lack of Ecological Validity of the Experimental Paradigms

With few exceptions (Nadelhoffer 2011), most philosophers have harshly criticised the validity of any theoretical conclusion drawn from neuroscientific experiments investigating intentional agency. Widely shared objections emphasise the lack of ecological validity of this kind of investigation, pointing out that the adopted experimental paradigms are unsound because ultimately unable to really tackle issues such as intentional agency and free will.

One suggested limitation depends on the type of free intentional action that experimental subjects are required to perform, which is deemed as unable to elicit feelings vaguely similar to a desire to act (Mele 2009). Such a limitation would make the corresponding results hardly generalisable to daily life scenarios in which individuals are in the position of making choices between options with appreciably different consequences. Given the absence of real reasons, of the type featuring in the causal theory, motivating them to act, participants have to be artificially *instructed* to make reportable unplanned *intentional* actions in a relatively constrained time window. In sharp contrast to this model of intentional action, Balaguer suggests that individuals are truly acting freely only when they make choices between options about which they are authentically torn, such as life-changing decisions about different job opportunities (Balaguer 2009). In contrast, the decision-making context associated to neuroscientific paradigms looks more similar to a situation where I am in front of a shelf in a supermarket and have to repeatedly choose between identical boxes of cereals - something I can do almost automatically or at least without the vigilant monitoring of conscious mental states. A related interpretative problem might arise if the type of actions experimental subjects make are so low-level that they become automatic or absent-minded: As a consequence, the data might be scarcely informative regarding the neural bases of conscious intentions. In this case, the absence of early conscious awareness might indicate that participants were not actually performing a full-fledged intentional action, but a merely voluntary action. Since many voluntary actions are performed in the absence of conscious intentions to act, some critics have plausibly argued that the absence of early conscious awareness *per se* might be considered irrelevant for denying free will (Lavazza & De Caro 2010; Nahmias 2014).

A second recurring criticism, also regarding the type of intentional actions featuring in neuroscientific experiments on voluntary processes, concerns the exclusive focus on *immediate* intentions. Many causal theorists indeed defend a two-tier account of intentions, whereby intentions concerning the immediate present are distinguishable from long-term intentions concerning the future. Examples of this distinction can be found in Searle (prior intention vs intention-in-action (Searle 1983)), Mele (distal vs proximal intention (Mele 1992)), and Bratman (future-directed intention vs present-directed intention (Bratman 1987)). For example, according to Bratman's influential theory of agency, it is the distinctively human capacity to entertain future-directed intentions that provides the necessary connection between motivation and deliberation (Bratman 2007). In a similar vein, Shaun Gallagher has suggested that higher order planning, and not single specific motor actions, may harbour free will (Gallagher 2006).

A third criticism relies on the fact that participants are mostly required, especially in experimental settings adopting Libet's paradigm, to report the time W, namely the timing at which they consciously experience they wanted to make a movement. This moment of awareness seemingly corresponds to a second-order state or meta-state (i.e., the consciousness of the wish to make a movement), rather than to a first-order intentional state (i.e., the wish to make a movement). It is nonetheless unclear whether this meta-state must have a specific causal role in order to support intentional action control. It seems indeed plausible to assume that the awareness of being in a given motivational state occurs later than the motivational state itself (Roskies 2010; see also Dennett & Kinsbourne 1992 for an early criticism about the temporal mismatch between the subjective experience of will and the perceived position of the clock hand).

By my lights, these criticisms target real deep issues in the experimental tradition initiated by Libet. However, whereas it is plausible to argue that Libet's experiments offer an impoverished, laboratory-based, representation of daily life choices, it would be inappropriate to claim that they do not offer an example of a type of intentional action people are in the position to perform. Indeed, despite the decision between identical boxes of cereals having no appreciable consequences, our intuition is that we are performing a free intentional (i.e., willed) action in grasping box A. The mere fact that we do not care about the possible consequences of an action does not turn an intentional action into an unintentional one. In this sense, one of the reasons why agency experiments mostly feature actions bearing no relevant consequences depends on how intentional actions are often conceptualised in cognitive neuroscience: Intentional or voluntary actions, as derived from specific motor circuits in the brain, originate from internally self-posited goals. In contrast, unintentional or involuntary actions, such as reflexes or involuntary movements, are triggered by external environmental stimuli or internal automatic states of the system (Passingham et al. 2010). Internally set goals may or may not depend on higher order reasons for acting, without this necessarily affecting the motor system. In principle, there is no clear cut, non-arbitrary, divide between high complex actions, as derived from conceptual reasoning, and simple spontaneous movements. To some extent, selecting between alternatives with no appreciable differences quintessentially exemplifies spontaneous decisions. If participants decide to move at *t1* (option 1) and not at *t2* (option 2) in a Libet's experiment, such a decision cannot be attributed to any immediately noticeable feature of the two options (external signal), but is seen as more likely to derive from purely endogenous processes.

Furthermore, in response to this type of objection, some recent experimental paradigms obtained comparable results (i.e., presence of RP for voluntary movements) while creatively bypassing the paradox of artificially prescribing people to perform intentional actions. For example, Khalighinejad et al. investigated the preparatory process (RP) prior to voluntary motor actions, by adapting a perceptual decision task in which participants were required to detect the motion of a display of dots towards the left or the right side of the screen. In each trial, participants had no clue about when the dots would have started moving coherently to the right or to the left. If participants did not wish to wait anymore, they had the option to press a *skip* button and move to the next trial. Voluntary actions were then operationalised as self-initiated skip responses while waiting for the display of dots to move coherently towards the left or the right. This way, the experimental paradigm was able to elicit the performance of intentional, or at least voluntary, actions without artificially probing participants to act (Khalighinejad et al. 2018; see also Murakami et al. 2014).

Finally, it is certainly plausible that, while walking through the aisles of the supermarket, I grasp a box of cereals remaining completely absent-minded. Such an absent-minded action can be easily intended as voluntary, if not intentional. It is indeed plausible that the lack of conscious awareness at the time of the action does not turn the action into an involuntary one. The very simple actions participants are required to perform in neuroscientific experiments seem to be of the type that might be executed completely absent-mindedly (thus without conscious awareness), remaining nonetheless voluntary. However, this assumption would require not to take into account the specific experimental setting, whereby participants are explicitly asked to pay attention to their conscious mental states in order to report them at a later stage. Despite being of the type that can be performed absent-mindedly, the choices typically made in neuroscience experiments are thus unlikely to be made absent-mindedly. Indeed, the corresponding experimental paradigms often explicitly require people to pay attention to their conscious intentions. Using Libet's temporal judgment task in an fMRI study, Lau et al. investigated the differences in the BOLD signal by comparing a condition where experimental subjects had to report the time at which they felt the intention to move *vs* a condition where they had to report the time at which they actually executed the movement. The data showed that, when participants were explicitly required to pay attention to their intentions, as it was the case in the classic Libet's task, the blood oxygenation level-dependent (BOLD) signal in the brain areas representing intentions (pre-SMA region of the medial prefrontal cortex) was enhanced (Lau et al. 2004. See also Haynes 2007). So, whereas conscious awareness might not be required for voluntary actions in general and for the type of actions performed in neuroscientific experiment in particular, its late (i.e., following the onset of the RP) appearance in cases where subjects were explicitly asked to pay attention to their intentions cannot be so easily dismissed as irrelevant.

5 Counter-arguments Based on the Lack of Causal Clarity

The RP has been interpreted as a reliable build-up of neuronal activity consistently preceding voluntary bodily movements, and thus specifically associated to motor preparation (Kornhuber & Deecke 1990). However, the potentially causal nature of the relationship between the RP, the time W, and the subsequent overt movement is still a matter of debate. Indeed, given that a clear causal linkage between the RP and the bodily movement cannot be established, the claim that bodily movements are caused by unconscious neural antecedents, and not by conscious mental states, seems to falter. In discussing whether the RP causes W, Haggard and Eimer provided evidence that the onset of the RP varies independently of W. In particular, trials where participants show an early W were characterised by a late onset of the RP, compared to trials characterised by a late W. A better candidate for *causing* W might be the *Lateralised readiness potential* (LRP), an increase in the electrical negativity in the area contralateral to the subsequent bodily movement, and reflecting the preparation of a specific movement after the action selection is made. Coherently, the LRP began earlier for actions with an early W, compared to actions associated to a late W. This important result suggests that the origin of W is likely to be related to a specific bodily movement rather than to a general preconscious motor preparation (Haggard and Eimer 1999).

For what concerns the relation between the RP and the overt bodily movement, an open possibility is that the RP plays a preparatory role without being sufficient for the movement to occur: In the absence of a causally relevant conscious decision to move, the subject might not execute the corresponding movement, despite the presence of the RP. In this light, Pockett and Purdy observe that the RP is neither necessary (single trials and even single experimental subjects may not show any RP) nor sufficient (changes in the electrical activity of the brain strictly resembling the RP waveforms do not necessarily originate movements) for action production (Pockett & Purdy 2011). Establishing a clear causal relation between the RP and the subsequent bodily movement implies at least the presence of the RP at the level of the single trial. By contrast, Libet's findings are based on averaging a number of trials where participants performed identical spontaneous bodily movements. In case the conscious decision to move is needed for the action to occur, conscious mental states might not have a role in action initiation, but might contribute (at least in full-fledged, non absent-minded intentional actions) to deciding whether to act or not. Libet himself, with a criticisable dualistic move, hinted at a similar solution, by arguing that, in the 150 ms preceding the onset of the movement, the subject could still decide to abort the action (Libet 1985).

By delving further into this possibility, some recent lines of research have been read as potentially reconciling the received view on intentional action with findings in the neuroscience of agency. In particular, Schurger et al. 2012 have questioned the widely accepted assumption that the RP can be truly recognised as the signal, within the motor system, of planning, preparing, and initiating a voluntary action. In contrast, they observe that changes in the neural activity preceding the bodily movement may merely reflect internal physiological noise, depending on spontaneous fluctuations and thus not specifically related to motor preparation. Within the proposed model, the timing at which the motor action occurs can be explained by means of an accumulator model: the bodily movement randomly occurs when the spontaneous electrical activity reaches a threshold, corresponding to the neural decision to move. The specific shape of the RP, normally detected before voluntary bodily movements, would be no more than an artefact due to the process of averaging the relevant epoched data time-locked to the onset of the bodily movement. Interpreted as averaged noise, the RP could hardly be seen as a signal of preconscious intentional motor preparation, playing rather a dispositional and predictive role. More speculatively, Schurger et al. suggest that the neural *decision* *to move now*, corresponding to the random crossing of the sensory-motor threshold, might be close in time to the timing of the awareness of wanting to make the movement (occurring about 150 ms before the action). The temporal overlapping between the neural commitment to move (the real decision) and the subjective awareness of wanting to make a movement might suggest that the received picture of intentional agency and the neuroscientific data are not so distant as it may seem.

The neuroscience of volition has devoted a considerable amount of research to identifying the point of no return, i.e., the timing after which a subject cannot prevent herself from acting (Libet 1985; De Jong et al. 1990). In a study looking at people's ability to veto their spontaneously initiated movement, Schultze-Kraft et al. (2016) successfully detected the presence of a signal like the RP in real time, at the level of the single trial. In each trial, subjects were instructed to make intentional unpredictable bodily movements unless they received a stop signal from the computer. In the first part of the study, a Brain-Computer Interface (BCI) was trained to recognise the typical RP for each subject and predict upcoming movements. In the second part of the study, the BCI predicted bodily movements in real time (by relying on neural antecedents) and sent a stop signal to the experimental subject. The results showed that, after the onset of a RP-like signal was detected, it was still possible for the subject to avoid moving in case the stop signal was sent earlier than 200 ms before EMG. This can be taken as evidence that the RP does not necessarily lead to the occurrence of the bodily movement, which can be still cancelled by the subject until a point of no return.

The results are compatible with those by Schurger et al. 2012 to the extent that they suggest that the real decision to move (or to abort the action) does not coincide with the onset of the RP but, more or less, with the timing of conscious awareness. Clearly, the point of no return and the neural decision to move are not expected to play the same causal role. While the neural decision to move might be a necessary causal component of an intentional action, the point of no return can be just a temporal threshold after which an unconsciously initiated movement cannot be aborted. However, this temporal association between the real commitment to move (i.e., neural decision to move and point of no return) and the awareness of wanting to make a movement has been hailed as having the potential to reconcile the neuroscientific findings on voluntary processes with more traditional views on intentional agency (Schurger et al. 2016).

The power to inhibit self-initiated actions is a key feature of human action control. Nonetheless, it is unclear whether the aforementioned findings provide a picture of agency that is truly compatible with the one framed by the causal theory. Indeed, these models do not make explicit assumptions about the origin of the neural commitment to move, supposedly following the onset of the RP and preceding the overt bodily movement. Clearly, the temporal closeness between the commitment to move and the time of conscious awareness *per se* cannot be interpreted as evidence of the fact that conscious mental states cause bodily movements: Also conscious decisions might originate from unconscious brain processes. In this light, recent experimental data suggest that the decision to inhibit or delay an action is similarly driven by antecedent and unconscious brain activity (Filevich et al. 2013).

6 A Plea for a More Constructive Role to Play for Neuroscientific Findings

Overall, most of the aforementioned discussion points originate from a similar underlying theoretical position, questioning whether neuroscientific experimental findings pose a real threat to the received view about the link between mental states and actions. Such a received view, which is well represented by the causal theory, sees conscious mental states as causing intentional actions, thus allowing agents to exercise intentional action control over their behaviour. In this respect, many argue that the neuroscientific findings that seemingly contradict the causal theory are too fragile to justify a rejection of the corresponding picture of intentional agency (Pereboom & Caruso 2018).

Whereas I also agree with this conclusion (i.e., there is not enough evidence that the causal theory is wrong), I claim that the problem with this argumentative strategy is that it implicitly assumes that specific philosophical claims (e.g. about agency or free will) can be directly (dis)proven by means of empirical investigation. According to this model, philosophical theorising is necessarily prior to empirical investigation, which has an essentially (dis)confirmatory role with respect to previously formulated philosophical claims. However, while the causal theory might work as a general framework for intentional agency (i.e., conscious mental states have a role in action production) and is clearly incompatible with other proposals (i.e., epiphenomenalism), single claims that might be inferred from the theory (i.e., how mental states should behave within the system) display little empirical translatability. As a result, when taken too literally, the causal theory may get things backwards. By contrast, my starting point is that the causal theory is meant to provide an artefactual model, and not a mechanistic explanation, of voluntary processes.

In my views, the results from cognitive neuroscience are not to be treated as a touchstone for philosophical claims, but can directly contribute to fuelling philosophical theories. The aspiration of the chapter is thus to suggest that philosophical analysis and cognitive neuroscience can actually work together, in a mutual exchange, to ameliorate our comprehension of intentional agency. This implies that, if needed, some features of intentional agency as framed by the causal theory must be set aside. To begin with, the understanding of intentional agency offered by the causal theory has little in common with the definition of intentional or voluntary processes at play in the empirical investigation. As previously mentioned, the former insists on the appropriateness (e.g. the end-state cannot be achieved through a deviant causal chain) of the causal link between intentional mental states and bodily actions. The latter more broadly distinguishes between internally generated actions, as derived from autonomously set goals, and externally prompted movements, elicited by stimuli present in the environment (Passingham et al. 2010). This theoretical incommensurability, beginning at the basic level of operational definitions, ramifies into more high level discussions about the role of conscious mental states (Mele 2010). However, in the next two subsections, I will advance some proposals and introduce a few examples about how results in cognitive neuroscience and philosophical analysis can jointly contribute to foster understanding of intentional agency.

6.1 *Multi-level Intentions*

As mentioned earlier in the chapter, the most radical theoretical conclusion one may draw from the empirical data on voluntary processes is some form of epiphenomenalism. In response to that, many have argued that the presence of unconscious neural antecedents does not prove that the conscious decision to move has no causal role. This claim might be justified by some of the pieces of evidence discussed above, both in a negative (e.g. no evidence of a direct causal connection between the onset of the RP, W, and the bodily movement) and in a positive (e.g. possible role of the neural decision to move in association to W) fashion. The emerging idea from research in cognitive neuroscience is that conscious mental states might not play a role at the time of action initiation, but could be crucially involved in other stages of action production, such as the decision to move now (Schurger et al. 2012), or the selection between abstract action alternatives (Rowe et al. 2000). This move already implies a departure from the causal theory, to the extent that action production is not seen as a substantially unitary process, firmly guided by a conscious intention to act that remains active from action planning to action execution (i.e., from my conscious abstract desire to drink water to my hand grasping that glass on the table).

Research on the cognitive architecture of the brain has shown that the process of selecting between alternatives, planning, and executing an action involves multilayer interacting structures. At the level of action selection between alternatives that are present in the environment, the integration of sensory information and internal representation of the state of the system is the basis for computing an action plan that is ultimately realised by muscular contraction. In the case of internally generated actions, the initial intention to act is abstract in the sense that it does not include all the details of motor execution, which are fixed by specific motor programs deriving from sub-actions (i.e., stages that are intermediate between abstract intentions and motor programs). The prefrontal cortex (PFC), and in particular the lateral prefrontal cortex (LPFC), has long been identified as the area of the brain where action generation and control take place (Duncan & Owen 2000; Miller & Cohen 2001). The proposed models of the functional organisation and cognitive architecture of the LPFC generally agree on the organisation of action control along an anterior-posterior axis, subserved by different sub-areas characterised by specific functions (Badre 2008; Bunge & Zelazo 2006; Koechlin et al. 2003; Koechlin & Summerfield 2007; Fuster 2004; Petrides 2005).

In particular, the *cascade model* proposed by Koechlin et al. 2003 and Koechlin & Summerfield (2007) describes action control as a hierarchically ordered process made possible by a cascade of top-down control from rostral to caudal LPFC and premotor regions, with anterior areas devoted to deliberative, abstract, temporally extended, action control (Grafton & Hamilton 2007; Haggard 2008; Hamilton & Grafton 2007; Kilner 2011). The model proposes that different areas of such a control network in the LPFC are responsible for executive control, defined as the capacity to select specific actions in relation to goals, thus resolving the entropy or competition between multiple action representations. Throughout this multilayer system, executive control and action coordination would nevertheless be guaranteed by the tight integration of information across the various specialised prefrontal regions. In fact, each stage in the hierarchical structure both exerts control over lower level representations and is controlled by the higher stages. These different sub-regions differentiate according to various degrees of flexibility and capacity of abstraction (i.e., their capacity to generalise across sets of representations) from the immediate action context.

Action control might thus be implemented by means of a complex hierarchical structure where different area jointly contribute to producing a given end-state. Within this framework, looking for a specific intentional state as able to play a specific causal role throughout the whole process of action production looks inevitably problematic. A more promising strategy consists in revisiting the structure of the causal theory, by accommodating this more complex view about how voluntary processes are causally integrated. For example, Pacherie has argued that goal-directed behaviour involves different levels of action specification, subserved by three types of intentions. Distal intentions (D-intention) operate at the highest, and more abstract, level, by setting up the overarching goal of an action and the appropriate sub-goals that are necessary to reach it. Proximal intentions (P-intentions) transfer the general action plan, set up at the previous level, into the current situation of action and select the appropriate motor planning. Finally motor intentions (M-intentions), corresponding to motor representations, are in charge of setting the finest parameters and values in order to execute the action, by using external sensory information. Crucially, Pacherie suggests that the content of motor intentions may not always be accessible to consciousness (Pacherie 2008, 2015).

6.2 *Disappearing Intentions*

Acknowledging the limitations of the causal theory, some philosophers have recently moved towards solutions that bypass the aforementioned theoretical criticisms (e.g. disappearing agent, deviant causal chain), by suggesting that intentional actions have to be treated as primitive, intrinsically unified, phenomena in our psychological ontology (Ford 2011; Levy 2013, 2015; O'Brien 2010). According to Levy, the reductive program advocated by the causal theory must be abandoned since intentional actions are not analysable in terms of primitive constitutive elements (i.e., intentional mental state and bodily movement). In contrast, intentional actions are to be conceived as basic. Levy defines intentional actions as bodily movements that people can stop and continue making intentionally. One of the benefits of this account is that it can be extended to voluntary, non-intentional, actions, such as instances of negligent behaviour, habitual actions, absent-minded behaviour, and possibly episodes of weakness of the will: If you absent-mindedly step on somebody's toes, you can then step back (Levy 2013). In comparison, bodily movements such as reflexes or spasms fail to satisfy the credentials for intentional or voluntary agency because you cannot intentionally decide to stop them. Abolishing the conjunctive causal representation of intentional action further favours the unified treatment of intentional bodily and mental actions - the latter being notoriously problematic within the causal theory. In fact, in the case of mental actions, a clear separation between causes and effects is hard to establish: How can I come to pay attention to a content without somehow already attending to that content? (Proust 2001).

Such a perspective is genuinely productive to the extent that it eliminates the strong dichotomy between different types of voluntary processes. General types of human voluntary behaviour lie along a continuum that goes from simple reflexes to higher complex functions (Haggard 2014), whereby the respective tokens are characterised by varying degrees of action control.

This view resonates with a scientifically reputable view of intentional agency that characterises voluntary processes in terms of freedom from immediacy (Gold and Shadlen 2007). Furthermore, a crucial role is played by the agent's ability to decide *whether* to act or not (or, as in Levy 2013, to continue to act or not) that has been clearly identified as a key element of action control (Haggard 2008). However, to avoid the pitfall of treating empirical findings as the test-bed for theoretical models, it should be specified that such a capacity to refrain from acting or stop acting, as the key feature of intentional or voluntary actions, is not to be taken in its literal sense. Indeed, as previously mentioned, we know from neuroscience that people are not able to prevent their intentional actions from occurring after a point of no return. So, paradoxically, people are acting intentionally even when they are not (anymore) in the position of preventing themselves from acting. The proposal of treating intentional actions as primitive bypasses causality (of mental states with respect to intentional actions) *tout court*: The unit for further analysis becomes the intentional action as a whole.

A related strategy consists in diminishing the centrality of intentions as discrete entities governing the physical body, without giving up on causality as such. The operation is similar to what suggested in the previous section to the extent that a partial departure from the causal theory might be required. Indeed, we might have to accommodate a wider notion of causation with little in common with the link between intentional states and actions articulated by the causal theory. In contrast with previously discussed hierarchical models of action control, Schurgher & Uithol 2015 and Uithol et al. 2014 have pursued a strategy of this sort. They suggest that the kind of information processing occurring at the level of the PFC during action control is ultimately incompatible with the representation of intentions as context-independent, inherently causal, discrete (in terms of content and functional role) entities. The authors argue in favour of a more dynamic and context-dependent model that does not fit a literal interpretation of the causal theory: The problematic step consists once again in moving abruptly from the theoretical model framed by the causal theory to the expectation that similarly contextual-free and discrete neural realisers are present in the brain.

In support of their position, the authors bring empirical evidence speaking against the thesis that context-free, discrete, states can sit at the top of the action-hierarchy. To mention one example, research by Fuster (2001) and Fuster et al. (1982, 2000) on single cell is interpreted as showing that even the most anterior parts of the LPFC (i.e., the area usually associated to high level, abstract, deliberative control) must rely on context-sensitivity to integrate information over time. Uithol et al. 2014 observe that distinguishing the contribution of different areas of the control network according to the different capacity for abstraction is misleading. Such a distinction is modelled on the gradient that goes from abstract intention (the intention to drink water) to concrete bodily action (grasping that glass on the table) framed by the causal theory. In contrast, the differences between areas of the LPFC is to be understood in terms of the type, source, and complexity of the information that are processed and integrated in order to produce a given end-state. In particular, more anterior areas integrate information pertaining to different types and sources (e.g. information from multiple senses), while caudal areas, which are devoted to low-level motor control, have to deal with specific information within the same type (e.g. effector-specific). Action control as a whole is thus realised through the integration of interdependent control processes in continuous relationship with contextual elements: Increasing complexity does not per se deny some form of mentally-related causality, but the kind of causality at play is substantially different from the one framed by the causal theory.

7 Conclusions

As human beings we are able to interact with the environment in various distinct ways, ranging from very simple motor actions to the implementation of long-term plans. In this chapter, I have argued that the major challenge for philosophy and cognitive neuroscience is to give reason to such a variety of instances of voluntary behaviour in a coherent manner. My overall goal was not to offer a coherent story of how intentional and voluntary agency unfold, but to provide some suggestions about possible fertile research pathways at the intersection between the philosophy and the neuroscience of action.

In philosophy, the causal theory frames intentional action control in terms of the causal authority exercised by conscious mental states over actions. Most philosophers have thus denied that findings in cognitive neuroscience can represent a real threat for the classic architecture of intentional agency as such. Whereas recognising that specific criticisms to experimental paradigms are appropriate, I have advocated a different perspective, whereby empirical findings are not necessarily to be treated as the touchstone for philosophical claims, but possibly contribute to building the theory itself. In this light, moving beyond rigid dichotomies, I suggested that some elements of the causal theory (e.g. discreteness of mental states, conscious access to all types of intentional states) might have to be abandoned in favour of a more articulated and nuanced understanding of voluntary processes.

References

Aguilar, J.H., & Buckareff, A.A., Eds. (2010). *Causing human actions. New perspectives on the causal theory of action*. MIT Press.

Anscombe, E. (1957). *Intention*. Basil Blackwell.

Arpaly, N., & Schroeder, T. (1999). Praise, blame, and the whole self. *Philosophical Studies*, 93(2):161–188.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5):193–200.

Balaguer, M. (2009). *Free will as an open scientific problem*. MIT Press.

Bratman, M.E. (1987). *Intention, plans, and practical reason*. Cambridge University Press.

Bratman, M.E. (2000). Reflection, planning, and temporally extended agency. *Philosophical Review*, 109(1):35–61.

Bratman, M.E. (2007). *Structures of agency: Essays*. Oxford University Press.

Bunge, S.A., & Zelazo, P.D. (2006). A brain-based account of the development of rule use in childhood. *Current Directions in Psychological Science*, 15(3):118–121.

Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60(23): 685–700.

Davidson, D. (1973). Freedom to Act. In: D. Davidson (1980), *Essays on actions and events*. Oxford University Press, 63–81.

De Jong, R., Coles, M.G., Logan, G.D., & Gratton, G. (1990). In search of the point of no return: The control of response processes. *Journal of Experimental Psychology: Human Perception and Performance*, 16(1):164–182.

Deecke, L. (2000). The Bereitschaftspotential as an electrophysiological tool for studying the cortical organization of human voluntary action. *Supplements to Clinical Neurophysiology*, 53:199–206.

Dennett, D.C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, 15(2):183–220.

Desmurget, M., Reilly, K.T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science*, 324(5928):811–813.

Doris, J.M. (2016). *Talking to our selves: Reflection, ignorance, and agency*. Oxford University Press.

Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. MIT Press.

Duncan, J. & Owen, A.M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, 23(10):475–483.

Filevich, E., Kühn, S., & Haggard, P. (2013). There is no free won't: Antecedent brain activity predicts decisions to inhibit. *PLoS One*, 8(2):e53053.

Fischer, J.M. (1999). The value of moral responsibility. *The Proceedings of the Twentieth World Congress of Philosophy*, 1:129–140.

Fischer, J.M., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge University Press.

Ford, A. (2011). Action and generality. In: A. Ford, J. Hornsby, F. Stoutland, Eds., *Essays on Anscombe's 'Intention'*. Harvard University Press, 76–104.

Fourneret, P., & Jeannerod, M. (1998). Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologi*a, 36(11):1133–1140.

Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1):5–20.

Frankfurt, H. (1977). The problem of action. *American Philosophical Quarterly*, 15(2):157–162.

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron*, 69(3): 548–562.

Fuster, J.M. (2001). The prefrontal cortex-An update: Time is of the essence. *Neuron*, 30(2):319–333.

Fuster, J.M. (2004). Upper processing stages of the perception-action cycle. *Trends in Cognitive Sciences*, 8(4):143–145.

Fuster, J.M., Bauer, R.H., & Jervey, J.P. (1982). Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks. *Experimental Neurology*, 77(3):679–694.

Fuster, J.M., Bodner, M., & Kroger, J.K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, 405(6784):347–351.

Gallagher, S. (2006). Where's the action? Epiphenomenalism and the problem of free will. In: W. Banks, S. Pockett, & S. Gallagher, Eds., *Does consciousness cause behavior? An investigation of the nature of volition*. MIT Press, 109–124.

Ginet, C. (2008). In defense of a non-causal account of reasons explanations. *The Journal of Ethics*, 12(3/4):229–237.

Gold, J.I., & Shadlen, M.N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30:535–574.

Grafton, S.T., & Hamilton, A.F. de C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4):590–616.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews: Neuroscience*, 9(12):934–946.

Haggard, P. (2014). Intention and agency. In: M.S. Gazzaniga, & G.R. Mangun, Eds., *The cognitive neurosciences, v*. MIT Press, 875–885.

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, 126(1):128–133.

Haji, I. (1998). *Moral appraisability: Puzzles, proposals, and perplexities*. Oxford University Press.

Hamilton, A.F. de C. & Grafton, S.T. (2007). The motor hierarchy: From kinematics to goals and intentions. In: P. Haggard, Y. Rossetti, M. Kawato, Eds., *Sensorimotor foundations of higher cognition attention and performance* (*Attention & performance*). Oxford University Press, 381–408.

Haynes, J.D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R.E. (2007). Reading hidden intentions in the human brain. *Current Biology*, 17(4):323–328.

Khalighinejad, N., Schurger, A., Desantis, A., Zmigrod, L., & Haggard, P. (2018). Precursor processes of human self-initiated action. *Neuroimage*, 165:35–47.

Kilner, J.M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences*, 15(8):352–357.

Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6):229–235.

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648):1181–1185.

Kornhuber, H.H., & Deecke, L. (1990). Readiness for movement - The Bereitschaftspotential-Story. *Current Contents Life Sciences*, 33(4):14.

Lagnado, D.A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, 108(3):754–770.

Lavazza, A., & De Caro, M. (2010). Not so fast. On some bold neuroscientific claims concerning human agency. *Neuroethics*, 3(1):23–41.

Lau, H.C., Rogers, R.D., Haggard, P., & Passingham, R.E. (2004). Attention to Intention. *Science*, 303(5661):1208–1210.

Levy, Y. (2013). Intentional action first. *The Australasian Journal of Philosophy*, 91(4):705–718.

Levy, Y. (2015). Action unified. *Philosophical Quarterly*, 66(262):65–83.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8(4):529–566.

Libet, B., Gleason, C.A., Wright, E.W., Jr., & Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (Readiness-Potential): The unconscious initiation of a freely voluntary act. *Brain*, 106(Pt 3):623–642.

Libet, B., Wright, E.W., Jr., & Gleason, C.A. (1982). Readiness-potentials preceding unrestricted "spontaneous" vs pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology*, 54(3):322–335.

Linser, K., & Goschke, T. (2007). Unconscious modulation of the conscious experience of voluntary control. *Cognition*, 104(3):459–475.

McKenna, M. (2011). Contemporary Compatibilism: Mesh Theories and Reasons-Responsive Theories. In: R. Kane, Ed., *Oxford Handbook of Free Will*, 2nd ed. Oxford University Press, 175–198.

Mele, A.R. (1992). *Springs of action*. Oxford University Press.

Mele, A.R. (2010). *Effective intentions. The power of conscious will.* Oxford University Press.

Mele, A., Ed. (1997). *The philosophy of action*. Oxford University Press.

Meltzoff, A.N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31(5):838–850

Miller, E.K., Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24:167–202.

Moore, J.W., Wegner, D.M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, 18(4):1056–1064.

Murakami, M., Vicente, M.I., Costa, G.M., & Mainen, Z.F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. *Nature Neuroscience*, 17(11):1574–1582.

Nadelhoffer, T.A. (2011). The threat of shrinking agency and free will disillusionism. In: W. Sinnott-Armstrong, L. Nadel, Eds., *Conscious will and responsibility: A tribute to Benjamin Libet*. Oxford University Press, 173–188.

Nahmias, E. (2014). Is free will an illusion? Confronting challenges from the modern mind sciences. In: W. Sinnott-Armstrong, Ed., *Moral Psychology, vol. 4: Freedom and Responsibility*. MIT Press, 1–26.

Nisbett, R.E., & Wilson, T.D. (1977). Telling More than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, 84(3):231–259.

O'Brien, L. (2010). *Self-Knowing Agents*. Oxford University Press.

O'Connor, T. (2009). Agent-causal power. In: T. Handfield, Ed., *Dispositions and causes*. Oxford University Press.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107(1):179–217.

Pacherie, E. (2011). Nonconceptual representations for action and the limits of intentional control. *Social Psychology*, 42(1):67–73.

Pacherie, E. (2015). Conscious intentions. The social creation myth. In: T. Metzinger, & J.M. Windt, Eds., *Open MIND*, 29(T). MIND Group.

Passingham, R.E., Bengtsson, S.L., & Lau, H.C. (2010). Medial frontal cortex: From self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences*, 14(1):16–21.

Pereboom, D., & Caruso, G.D. (2018). Hard-incompatibilist existentialism: Neuroscience, punishment, and meaning in life. In: G.D. Caruso, & O. Flanagan, Eds., *Neuroexistentialism: Meaning, morals, and purpose in the age of neuroscience*. Oxford University Press, 193–222.

Petrides, M. (2005). The rostral-caudal axis of cognitive control within the lateral frontal cortex. In: S. Dehaene, J.-R. Duhamel, M.D. Hauser, & G. Rizzolatti, Eds., *From monkey brain to human brain. A Fyssen foundation symposium*. MIT Press, 293–314.

Pockett, S., Purdy, S. (2010). Are voluntary movements initiated preconsciously? The relationships between readiness potentials, urges and decisions. In: W. Sinnott-Armstrong, L. Nadel, Eds., *Conscious will and responsibility: a tribute to Benjamin Libet*. Oxford University Press, 34–46.

Proust, J. (2001). A Plea for Mental Acts. *Synthese*, 129(1):105–128.

Roskies, A. (2010). How does neuroscience affect our conception of volition? *Annual review of neuroscience*, 33:109–130.

Rowe, J.B., Toni, I., Josephs, O., Frackowiak, R.S., & Passingham, R.E. (2000). The prefrontal cortex: response selection or maintenance within working memory? *Science*, 288(5471):1656–1660.

Schlosser, M.E. (2010). Bending it like Beckham: Movement, control and deviant causal chains. *Analysis*, 70(2):299–303.

Schultze-Kraft, M., Birman, D., Rusconi, M., Allefeld, C., Görgen, K., Dähne, S., Blankertz, B., & Haynes, J.D. (2016). The point of no return in vetoing self-initiated movements. *Proceedings of the National Academy of Sciences USA*, 113(4):1080–1085.

Schurger, A., & Uithol, S. (2015). Nowhere and everywhere: the causal origin of voluntary action. *Review of Philosophy and Psychology*, 6(4):761–778.

Schurger, A., Mylopoulos, M., & Rosenthal, D. (2016). Neural antecedents of spontaneous voluntary movement: A new perspective. *Trends in Cognitive Sciences*, 20(2):77–79.

Schurger, A., Sitta, J.D., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences USA*, 109(42):E2904–E2913.

Searle, J. (1983). *Intentionality*. Oxford University Press.

Sehon, S.C. (2016). *Free will and action explanation. A non-causal, compatibilist account*. Oxford University Press.

Setiya, K. (2007). *Reasons without rationalism*. Princeton University Press.

Soon, C.S., He, H.A., Bode, S., & Haynes, J.D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences USA*, 110(15):6217–6222.

Soon, C.S., Brass, M., Heinze, H.J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11(5):543–545.

Strawson, P.F. (1962). Freedom and resentment. *Proceedings of the British Academy*, 48:1–25.

Uithol, S. Burnston, D.C., & Haselager, P. (2014). Why we may not find intentions in the brain. *Neuropsychologia*, 56:129–139.

Velleman, J.D. (1992). What happens when someone acts. *Mind, New Series*, 101(403):461–481.

Wegner, D.M. (2002). *The illusion of conscious will*. MIT Press.