



Annual Review of Psychology

The Neurocognitive Bases of Human Volition

Patrick Haggard^{1,2,3}

¹Institute of Cognitive Neuroscience, University College London, London WC1E 6BT, United Kingdom; email: p.haggard@ucl.ac.uk

²Institute of Philosophy, School of Advanced Study, University of London, London WC1E 7HU, United Kingdom

³Laboratoire de Neurosciences Cognitives, Département d'Études Cognitives, École Normale Supérieure, 75005 Paris, France

Annu. Rev. Psychol. 2019. 70:17.1–17.20

The *Annual Review of Psychology* is online at
psych.annualreviews.org

<https://doi.org/10.1146/annurev-psych-010418-103348>

Copyright © 2019 by Annual Reviews.
All rights reserved

Keywords

volition, action, will, human, neuropsychology

Abstract

Volition refers to a capacity for endogenous action, particularly goal-directed endogenous action, shared by humans and some other animals. It has long been controversial whether a specific set of cognitive processes for volition exist in the human brain, and much scientific thinking on the topic continues to revolve around traditional metaphysical debates about free will. At its origins, scientific psychology had a strong engagement with volition. This was followed by a period of disenchantment, or even outright hostility, during the second half of the twentieth century. In this review, I aim to reinvigorate the scientific approach to volition by, first, proposing a range of different features that constitute a new, neurocognitively realistic working definition of volition. I then focus on three core features of human volition: its generativity (the capacity to trigger actions), its subjectivity (the conscious experiences associated with initiating voluntary actions), and its teleology (the goal-directed quality of some voluntary actions). I conclude that volition is a neurocognitive process of enormous societal importance and susceptible to scientific investigation.



Review in Advance first posted
on August 20, 2018. (Changes may
still occur before final publication.)

Contents

1. INTRODUCTION AND SCOPE	17.2
2. DEFINING THE ELEMENTS OF VOLITION	17.3
2.1. Internal Generation: Volitional = Not Externally Triggered	17.3
2.2. Decisional Control: Volition = Not Habitual, Not Automatic	17.4
2.3. Goal Directedness: Reasons, Values, Outcomes	17.5
2.4. Spontaneity	17.5
2.5. Subjective Experience	17.5
3. METHODS FOR THE EXPERIMENTAL STUDY OF VOLITION	17.6
4. MECHANISMS OF VOLITION IN THE HUMAN BRAIN	17.7
5. GENERATIVITY IN HUMAN VOLITIONAL ACTION	17.9
6. THE SUBJECTIVE EXPERIENCE OF VOLITION	17.11
7. TELEOLOGY IN HUMAN VOLITION	17.13
8. WIDER SIGNIFICANCE: VOLITION AND RESPONSIBILITY	17.15
9. CONCLUSION	17.17

1. INTRODUCTION AND SCOPE

We normally have an effortless recognition of our own volition, in that we readily recognize an action as either what we intended to do or not. Moreover, our capacity for voluntary action is of great importance to us: We regret or complain vigorously when our capacity for voluntary action is reduced, for example, by disease or by constraints such as incarceration. Nevertheless, it is not easy to define volition, nor to pinpoint the difference between an action that is voluntary and one that is not. As discussed below, volition seems to involve several rather distinct aspects of action. In this review, I first outline these different elements, then discuss the distinctive cognitive and neural mechanisms that underlie volition, and finally consider the impact that psychological theories of volition might have on wider society. Throughout, I am interested only in volition as it relates to initiation and control of motor action. There are other, more psychic targets for volition: For example, can one have volitional control over one's thoughts, emotions, and other mental states? Volitional control of action may be more scientifically tractable than purely psychic volition because the former ends with an objective, measurable, physical event, while the latter does not.

I do not address the important issue of volitional inhibition of action. The capacity to inhibit an action, particularly a prepotent or impulsive action, is often considered a key marker of voluntary self-control (Mischel et al. 1972). In a weak sense, the capacity to inhibit makes all actions volitional. If someone is able to inhibit their actions, then, it is argued, they could have done otherwise when they do act. "Could have done otherwise" counterfactuals have been important criteria for voluntariness in philosophy. However, this criterion is harder to accommodate in behavioral science, which typically begins with the observation that a person did act, and cannot easily accommodate counterfactual descriptions. I therefore distinguish between positive volition, meaning the initiation of intentional action, and negative volition, meaning the capacity to inhibit actions that are inappropriate. The literature on neurocognitive processing of negative volition is somewhat separate from that on positive volition and raises a specific set of methodological difficulties (for a review, see Filevich et al. 2012).

This review therefore focuses primarily on volitional action initiation rather than on self-control through action inhibition. The main goal is to characterize the features, cognitive

17.2 Haggard

Review in Advance first posted
on August 20, 2018. (Changes may
still occur before final publication.)



Table 1 Some key features of volition and the possible constraints that they imply for the neuroanatomical substrate

Key feature of volition	Neuroanatomical constraint
Leads to movement	Strong connections to motor areas
No external trigger	Weak connections from sensory areas
Reasons responsive	Strong connections from valence and reward circuits
Outcome or goal directed	Strong connections with planning and monitoring areas
Spontaneous or innovative	Independence from subcortical habitual circuits
Involves consciousness	Connections to frontal and parietal cortices (remains controversial)

processes, and neural circuits that underpin volitional actions. This may contribute to the wider psychological project of a substantive explanation of the internal, intentional generativity that characterizes human behavior. To achieve this goal, I first investigate the defining features that distinguish volitional actions from other, nonvolitional actions. I then briefly outline the evidence that two broad cortical regions, the medial frontal and the parietal cortices, play a distinctive role in human volition. Next, I consider in detail three features that are particularly important to human volition, namely generativity, subjectivity, and teleology or goal directedness. I end with brief remarks about the significance of volition for responsibility.

2. DEFINING THE ELEMENTS OF VOLITION

Defining volition is difficult. **Table 1** contains some of the key features of volition, based on a range of philosophical, scientific, and other accounts. The striking feature of the table is the number of different, sometimes contradictory, connotations of the term. Given this heterogeneity, one might wonder whether there is a single mental factor of volition at all. One might also be surprised at how rarely people get confused in everyday life about whether a given action, either their own or another's, is volitional. For clarity, I briefly describe some of these features of volition and explain their background. I start from the view that the distinctive task of psychology is to explain behavior. "Why did she do that?" is the core question of psychology. The features described in **Table 1** constitute possible answers to this question. Given this starting point, it seems axiomatic that volition is a mental state that aims at movement, as opposed to some other biological or psychological function. Therefore, as explained above, I do not discuss the interesting topic of volitional control of purely mental contents, such as memories, images, and emotions.

Table 1 also shows that each candidate feature of volition has implications for how volition might be realized in the human brain. These factors are suggestions that may not be universally accepted or may not be decidable from current neural data. They are included to show that the debate over whether volition exists, and whether it is a genuine cognitive state or process amenable to scientific psychological investigation, can also be informed by neural data.

No single one of the features in **Table 1** seems necessary on its own to make an action volitional. However, several of them together may be sufficient. In this review, I try to go beyond the traditional definition of volition by exclusion to describe the neurocognitive features of volitional initiation of action. First, however, I describe each feature in turn.

2.1. Internal Generation: Volitional = Not Externally Triggered

Several authors have viewed volition as a matter of internal generation, or endogenicity, of actions. Internally generated actions are typically contrasted to reactive situations, in which movement is

a response to external environmental stimuli (Passingham et al. 2010). This approach identifies either two categorically distinct routes to action (Passingham 1987) or a continuous dimension with multiple intermediate possibilities between these extremes. One pole involves externally triggered responses, with simple reflexes as one obvious example. The opposite pole is sometimes defined by exclusion: an action that is not a reflex. Yet definition by exclusion adds little positive information about what volition is. In one view, volition may involve internally generated information regarding whether to make an action, what action to make, or when to make it (Brass & Haggard 2008).

Alternatively, internal generation has sometimes been defined as a case where the change that causes the action is inside the person (or animal), rather than in the external environment. This definition may also seem unsatisfactory because both internal and external events can trigger actions which seem reflex-like or compulsive. Yet somehow this internal triggering does not capture what we mean by volitional. For example, a sudden pain in my abdomen will cause me to double up and grasp the painful region: The action seems to be internally generated because it is triggered by some event in some internal organ. However, one might not wish to say that the action is volitional. Similarly, I may feel compelled to reach for a glass of water because I feel extremely thirsty. Some internal causes, such as decisions, may be associated with volition, whereas other internal events seem to drive responses more directly.

If volitional actions are not caused by external stimuli, then they must be caused by something else. Modern scientific thought eschews dualist accounts that posit a soul or homunculus as the cause of action. In fact, the distinction between reflex and internal generation may simply be a matter of directness of causal chains. In simple reflexes, a single, identifiable stimulus causes an immediate and spatially focused response. In other cases, an action can reflect an integration of many sources of input information over an extended time period. Consider voting in an election. This seems highly volitional (at least, it should be volitional). It is not independent of the external environment, but rather depends on an astonishing range of external factors, such as the candidate's previous actions, the party's manifesto, and the position of the candidate's name on the ballot paper. Volitional actions integrate across wide contexts, so it becomes hard to identify a single specific external trigger. Internally generated action does not mean that the action occurs out of nothing, but rather that it is related to many things (Schüür & Haggard 2011). In the words of Gold & Shadlen (2007, p. 542), such actions are "not governed by the immediacy of sensory input or motor output." This freedom arises because of the range and complexity of the information that is integrated to cause the action.

2.2. Decisional Control: Volition = Not Habitual, Not Automatic

Another approach to volition is based on contrasting volitional actions with automatic or habitual actions. Habitual actions are characterized by repeated, automatic triggering and stereotyped execution, with minimal cognitive involvement: "Why did she do that?" "Well, she just always does" (Dolan & Dayan 2013). This is, again, a definition by exclusion: An action is volitional to the extent that it does not occur stereotypically, routinely, and automatically. In this contrast, what makes an act volitional is its origin in a specific decision that this action is appropriate in the current context, and that the action is initiated based on current needs and goals. Whereas habitual actions are performed in the absence of any positive decision, I will perform a volitional action only when and because I decide to do so, and I will control its initiation accordingly.

This formulation captures the counterfactual, or "could have done otherwise," aspect to volitional action. A volitional action implies a choice, i.e., a positive initiating event or decision, in a way that a habitual action does not. This may be the choice of whether to act, the choice of which of several actions to make, or simply the choice between acting now and acting later. The choice

amounts to generating information that leads to specific details of action, such that a particular action would occur in some situations (possible worlds in philosophy terminology) but not in others.

2.3. Goal Directedness: Reasons, Values, Outcomes

Volitional actions are often considered to be actions made for a reason. The reason is typically a representation of some desirable goal state that is achieved or advanced through the action. This view has dominated the recent philosophy of action (Davidson 1963). For an action to be volitional, one should be able to ask a question of the form, “Why did she do that?” In general, the answer to this question should indicate a specific reason in the agent’s mind. This raises the central question of whether the agent is necessarily conscious of the reasons for their volitional action, or whether one can sometimes act volitionally as a result of unconscious reasons. A detailed discussion of this question would require a psychodynamic account and is beyond the scope of this article (Levy 2011). However, most thinkers agree that a greater degree of conscious awareness of the reasons for action corresponds to stronger volition.

Defining volition as goal directedness has clear implications for underlying cognitive mechanisms. At the very least, the agent must have a representation of a goal, must attribute net positive value to the goal, and must realize that they are not currently in the goal state but could attain this state through their own action. These cognitive processes would then generate an action that aims to bring the goal state closer. These cognitive processes imply a degree of counterfactual thinking, quite distinct from the representation of the current environment considered in the previous section (for a dissenting view, see Barandiaran et al. 2009). They also require acquiring, storing, and then recalling action–outcome associations so as to initiate the correct action given the desired goal.

2.4. Spontaneity

Sometimes, people and animals also generate salient, nonhabitual actions without apparent reason. Such spontaneous actions might reflect either an inevitable but undesired result of motor noise or a neural generator that occasionally emits exploratory actions (patch leaving in the terminology of animal foraging). This aspect of volition typically emphasizes an apparently aimless generativity of action, without reasons and without specifically intended reward: “I don’t know why, I just did it!” Of course, spontaneous, even random, generativity can sometimes be valuable. First, it can generate an action even when value-based decision making cannot, as in the thought experiment of Buridan’s ass, who starved through indecision when faced with two bales of high that seemed equally attractive. Second, foraging has epistemic value (Friston et al. 2015). Coupled with outcome-based learning, spontaneous action generation can provide an important method for discovering new and potentially valuable actions. In the psychological literature on voluntary action, this concept of spontaneous action applies to experimental studies that invoke an urge to act in the absence of any particular reason or reward for doing so (Libet et al. 1983), and also to some notions of creative actions.

2.5. Subjective Experience

A final aspect of volition is its relation to conscious experience. For an act to be volitional, one must be aware that one is acting and aware of initiating one’s act. For example, someone who sleepwalks is unaware of their walking action and did not choose to initiate it. Someone who makes a reflex



action in response to an immediate stimulus, such as a sneeze, is aware of their action but is also aware that they did not choose to initiate it. In general, volitional actions feel entirely different than physically comparable involuntary movements. Wittgenstein (1953, § 621) famously asked, “What is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?” He asked the question ironically, but one plausible answer to his question would be: a subjective experience of volition.

Is consciousness strictly necessary for volition? The traditional answer to this question is yes: An action can only be voluntary if the agent is aware of what they are doing. However, what precisely are they aware of? What is the content of the conscious part of volition? This question remains controversial. Philosophical theories, following Davidson (1963), generally insist that voluntary actions arise from intentions, which are typically considered to be conscious propositions or thoughts. In contrast, modern neuroscience views actions as the result of a causal cascade within the voluntary motor pathway. Some neural activities within this pathway may additionally cause conscious experience. In this view, the distinctive conscious experience of volition is a consequence of neural activity, rather than a cause of neural activity.

3. METHODS FOR THE EXPERIMENTAL STUDY OF VOLITION

Studies of effects of brain damage and disease in humans and other animals have contributed to understanding volition. Thus, focal lesions of the medial frontal cortex lead to a reduced level of spontaneous movement in primates (Passingham 1993). Interestingly, damage to the same general area in humans leads to an excessive degree of stimulus-triggered action, particularly if transcallosal fibers were also damaged (Boccardi et al. 2002, Della Sala et al. 1991). Neuropsychiatric conditions, notably psychoses, can lead to florid disturbances of volition, in which patients explicitly report that their actions are not under their own control (Frith 2014). The human neuropsychological literature on volition generally begins with an intuitive concept of normal voluntary action and relies on observation to identify how damage and disease alter the normal pattern. Some researchers have used experimental paradigms to identify how volitional actions of neurological (Kritikos et al. 2005) or psychiatric (Voss et al. 2017) patients differ from those of healthy volunteers. Nevertheless, explicit operational definitions of volition are rare in the neuropsychological literature. Instead, the value of this literature lies in showing that our somewhat intuitive concept of volitional control of behavior involves specific cognitive functions that have identifiable bases in the brain. There are relatively few cases of designed experiments with such patients, or of neuropsychological double dissociations within the broad area of volition. Therefore, neuropsychology has made only modest contributions to analyzing or fractionating volition into specific cognitive components (for an interesting exception in the case of inhibition of voluntary action, see Sumner et al. 2007).

Good experimental design in psychology treats the brain or mind as a black box. Manipulating the input (stimulus) and measuring output (behavior) allows one to understand the mental function. However, in volition, the input is, by definition, up to the participant themselves. If the experimental design is based on a controlled input, then volition seem to disappear. Moreover, the behavioral output is not easy to reverse engineer: Any given motor action might be caused by a volitional process or by an involuntary process. The physical features of movement are largely silent about the movement’s origins. For example, if an action is indeed volitional, then the agent might have initiated it for any of several reasons. This is why the psychologist’s question, “Why did she do that?” is so interesting and worth asking. In much of experimental psychology, the experimental conditions constrain the situation so only one answer to the question is possible: “Why did she press the button?” “Because the experimenter presented a Go stimulus.” However,

this does not work straightforwardly for studies of volition. If the experimenter brings about a volitional action by the normal methods of instruction, stimulus, and context, then the action seems to be externally triggered rather than truly volitional.

An alternative experimental tradition, going back to the Wurzburg school (Ach 1905), has resorted to instructions such as “Make an action whenever you feel like it” (e.g., Libet et al. 1983). This means that the input to the black box, and the reason for the initiation of action, are unknown, perhaps even unknowable. Importantly, this shifts the emphasis away from the reasons-responsive aspect of volition and toward the spontaneous, generative aspect. Some authors have argued that volition then becomes a mere artifact of the social situation, in which the participant behaves in the way that they believe will satisfy the experimenter’s notion of volition.

Because defining volition is not straightforward, experimental studies are often criticized for failing to focus clearly on volition. This critique can be either deflationary, i.e., the study is just about decision making, or inflationary, i.e., the study does not capture what is really meant by volition. To respond to both critiques, experimental designs can use an operational definition of volition by designing conditions that differ with respect to one or more of the features shown in **Table 1**. Thus, the most common experimental design for studying volition involves comparing conditions in which the details of an action (such as what action to make, when to make it, or indeed whether to act at all) are instructed by an external stimulus with one in which the participant freely chooses and generates these details for themselves (Brass & Haggard 2008, Deiber et al. 1991, Duverne & Koechlin 2017, Frith et al. 1991, Zapparoli et al. 2018). This contrast between free and instructed actions may not capture everything that one would wish to find in a theory of volition, but it does provide an operational definition of a factor that is clearly relevant to volition.

4. MECHANISMS OF VOLITION IN THE HUMAN BRAIN

Volition has been defined (**Table 1**) as a state that leads to movement. Therefore, the brain basis of volition should be identifiable by reverse engineering the motor pathways and tracing backward from muscle, to motoneuron, to descending command. This approach would involve following the final common path, in the words of Sherrington (1906), to find the original motor command for self-initiated movement.

This approach shows that the primary motor cortex (MI) is not a center for volition, although it clearly lies on the voluntary motor pathway. Artificial stimulation of MI produces involuntary movements that feel entirely different from voluntary actions (Haggard et al. 2002). Moreover, MI is involved in reflexive (Johansson et al. 1994) and involuntary movements (Ghosh et al. 2014). Two major frontal cortical areas send afferents to MI. These are the supplementary motor area (SMA),¹ located primarily in the medial wall, and the premotor cortex (PMC), located on the lateral surface of the hemisphere, immediately in front of MI. Both areas are implicated in the complex control of movement, and both exert a major influence on movement indirectly via MI, as well as by some direct projections to the cord. However, a striking double dissociation exists between the functional properties of these areas, which is highly relevant to the definition of volition given above. The SMA appears to provide neural drive to MI for self-initiated actions, while the PMC provides drive to MI for actions guided by external stimuli. For example, lesions of the SMA reduced monkey’s spontaneous movements, while leaving the capacity to respond conditionally to different visual stimuli intact. Lesions of the PMC had the opposite effects (Passingham 1987).

¹The historical SMA is now classically divided into two areas, the more anterior pre-SMA and the posterior SMA-proper, based on different anatomical connectivity (for a review, see Picard & Strick 1996).



Halsband and colleagues (1994) recorded from single neurons in both areas while monkeys made actions either internally-generated from memory, or triggered by external visual stimuli. They found that neurons in medial areas, such as the SMA, typically responded before or during internally generated movements, while those in the PMC responded predominantly before or during externally triggered movements (Halsband et al. 1994). Recordings of single neurons in the SMA and pre-SMA in humans undergoing neurosurgical evaluation often showed ramp-like changes in firing rate prior to the moment of self-paced actions in an experiment like that of Libet and colleagues (1983) (see Fried et al. 2011).

The results of stimulation studies broadly concur with the idea that the medial frontal cortex plays a special role in initiation of internally generated action. Direct stimulation of the SMA and pre-SMA areas in neurosurgical procedures sometimes elicits an experience described as an urge to move (Fried et al. 1991) without evoking actual movement. Interestingly, this experience is somatotopically specific, in that the patient can or does report to which specific body part the urge relates. This finding provides a crucial argument against the widespread skeptical claim that the experience of volition is merely a post hoc rationalization triggered by the need to make sense of one's own body movements (Banks & Isham 2009, Wegner 2002). Rather, premovement activity in these medial frontal areas would produce a conscious motor experience of volition, in somewhat the same way that neural activity in visual motion areas produces conscious visual experiences of motion. Interestingly, stronger stimulating current at the same locations would often evoke movements, typically of the same body part that was previously named in the report of urge to move. This could mean that the centers responsible for the experience of volition lie on the main voluntary motor pathway, which is activated artificially by direct stimulation. However, the result could also be explained by simple spread of current to the adjacent motor cortex through nonphysiological paths.

Interestingly, a recent study has reported a rather different experience of will to persevere during stimulation of the anterior mid-cingulate cortex, lying deeper within the medial wall, below the pre-SMA area. An illustrative patient report describes the feeling as: "push harder, push harder, push harder to try and get through this. . ." These experiences again occurred without evoked movement, although they were accompanied by autonomic responses. Interestingly, they included a strong emotional component apparently absent from SMA or pre-SMA stimulation (Parvizi et al. 2013).

An important current controversy involves the relative contributions of medial frontal versus parietal regions in volition. Evidence from psychophysical studies in patients with parietal lesions confirms an abnormal and delayed experience of volition in the Libet task (Sirigu et al. 2004). Again, direct stimulation in neurosurgical patients has emerged as a key experimental method, since it has allowed scientists to break the normal association between the subjective feeling of volition and the occurrence of voluntary movement. Desmurget and colleagues (2009) reported several cases of self-reported urge to move ("je veux bouger") during intraoperative stimulation of the parietal cortex. These were somatotopically specific and were not accompanied by evoked movement. Unfortunately, the setting of neurosurgical evaluation does not easily allow the kind of exploratory testing or detailed psychophysical and psychometric evaluations that the experimental psychologist might wish for. Therefore, it remains unclear how similar the experiences of volition elicited by medial frontal and parietal stimulation may be.

Given the evidence to date, it seems likely that both medial frontal and parietal areas contribute to the generation and subjective experience of voluntary action. However, the major anatomical projections of the two areas differ. This review begins the search for mechanisms of volition by reverse engineering the voluntary motor pathway. A major input to the frontal and prefrontal cortices comes from the basal ganglia via the thalamus. This ascending input is organized into

a series of parallel loops, which are thought to drive different forms of cognition and action (Alexander et al. 1986). Medial frontal areas associated with volition receive both direct thalamic inputs and indirect inputs relayed from other prefrontal cortical areas (Picard & Strick 1996). In contrast, the parietal cortex receives input primarily from posterior sensory areas. It sends output to the motor cortex for the guidance of movement and is also thought to receive direct or indirect reciprocal connections from motor areas in the form of efference copies (Desmurget et al. 2009). The parietal cortex is argued to generate conscious experience when it integrates these incoming efference copies. Furthermore, the content of the experience corresponds to the motor command information contained in the efference copy.

5. GENERATIVITY IN HUMAN VOLITIONAL ACTION

I define volition first and foremost as a state or process that leads to action. Historically, the study of human volition has been strongly linked to studies of the neural origin or generation of actions. The initial breakthrough came when Kornhuber & Deecke (1965) used back-averaging of human electroencephalogram (EEG) data time-locked to self-paced movements. Using this method, they found a gradual, ramp-like increase in negativity prior to volitional actions. This signal is not normally visible on any single trial because it is swamped by the much larger variation in background EEG. However, it can be extracted from noise by averaging across several movement epochs. The morphology of this potential recalled a build-up or preparation for movement, which may explain Kornhuber & Deecke's choice of name: *Bereitschaftspotential*, or readiness potential (RP). The RP often begins 1 s or more before movement onset and continues to rise until just before the onset of muscle contraction, at which point there is a sharp reversal into positivity. The potential is initially maximal over the frontal midline; it has been localized to the medial frontal cortex but is broadly distributed. The morphology of different subcomponents within the RP has been debated, but the most striking feature is a shift in the topography toward the hemisphere contralateral to the hand that will move. This lateralized readiness potential (LRP) occurs some 500–200 ms before the onset of the movement itself (Haggard & Eimer 1999). LRP onset is considered an important marker for studies of volition because it represents a time point by which information about which of the two hands will be used to make the action must have been generated. Neuroimaging studies suggest that it may correspond to a progression of activation from medial frontal regions into MI itself. In conditions where there is no external instruction regarding which hand to use, action selection must be internally generated. The LRP onset provides a chronometric marker that such selection has occurred. In practice, few paradigms have used LRP onset to investigate internal selection of which hand to use for volitional actions. One possible reason for this may be the difficulty of ensuring that participants make these decisions in real time, rather than predeciding a sequence of responses in advance (see Misirlisoy & Haggard 2014).

Importantly, the RP seems to mirror some of the key criteria for volition. Notably, the RP was found to be present prior to internally generated actions but absent prior to actions that were triggered by an unpredictably occurring external stimulus (Jahanshahi et al. 1995). Interestingly, if the external cue occurs at a predictable time (for example, at a fixed latency after a warning signal), then a potential known as contingent negative variation (CNV) occurs in anticipation of the external stimulus. The CNV and RP have similar topographies, and there has been extensive controversy regarding whether they represent two distinct brain processes or only one (Rohrbaugh & Gaillard 1983). This issue raises questions regarding the RP: Should it be considered as preparation or energization for movement, or as a passive anticipation, along the lines of a CNV, that an event is about to occur? The question remains unresolved.



Several studies have considered how different kinds of internal generative processes might affect the RP. Libet et al. (1983) made an interesting distinction between Type I and Type II RPs. Type I RPs featured an early onset, often 1.5 s or more prior to action, and a high peak amplitude. They were associated with an explicit conscious deliberation and with preplanning or preparation of the upcoming movement. In contrast, Type II RPs had a later onset, typically between 500 and 1,000 ms prior to action, and a lower maximum amplitude. They were associated with movements that occurred more spontaneously, without prior deliberation, but “freely capricious in origin” (Libet et al. 1983, p. 625). It remains unclear from Libet and colleagues’ description whether the distinction reflects an instruction that they gave to their participants or a commentary that they offered on their own experiences. This may explain why, despite the paper being highly cited, the Type I/Type II distinction appears not to have been taken up by the field. In any case, Libet and colleagues’ result does suggest that the RP is positively associated with cognitive engagement and effort with respect to the impending movement. The more the participant thinks about the action, the earlier and larger is the RP.

A similar conclusion follows from a more recent study (Verleger et al. 2016). Participants were simply instructed to make self-paced actions at a given rate, either explicitly counting the seconds between actions or not. Longer interaction intervals were associated with earlier and larger RPs, suggesting that the time available for the preparatory build-up may have a major influence on RP.

Taken as a whole, the RP literature has generally accepted that the RP is a readout of a distinctive cognitive process related to the volitional generation of action. This entire concept has been challenged by a recent study of internally generated action. Schurger and colleagues (2012) began with the well-established concept of accumulation of evidence in decision making, as implemented in drift diffusion models (e.g., Gold & Shadlen 2007). In such models, a decision and a motor response is triggered when accumulating evidence about a stimulus reaches a preset threshold level. Schurger and colleagues then hypothesized that the timing of internally generated actions, in the absence of any external stimulus, might depend on the accumulation by a leaky integrator mechanism of some unknown stochastic source of internal neural noise. On this view, the RP is essentially a readout (albeit an imperfect one) of the trajectory by which the random walk of neural activity is accumulated and approaches the threshold to trigger an action. Each individual action is preceded by a different pattern of random walk, but the approach to the threshold is always in the same direction, from below, since the threshold-crossing is the mechanism for triggering the action. When these different preaction signals are averaged across trials, they produce a gradual, ramp-like pattern with the same appearance as the RP. However, interpreting this pattern as the signature of a specific neural signal would be a mistake—indeed, there is no signal at all that triggers the action at a specific moment. Rather, the timing of actions is random, and the RP is an artifact of biased sampling of a noisy neural signal, with the bias coming from the fact that the extraction of the signal is locked to the threshold-crossing event that triggers the action. Schurger and colleagues’ model also makes clear predictions about the relationship between the form of the RP and the delay prior to self-initiated action. These predictions were experimentally confirmed (Schurger 2018). A simple stochastic model would also predict that the RP should have a fixed maximal amplitude, reflecting the height of the decision threshold above the initial level of the random signal. This prediction has not been systematically tested.

This model clearly shifts the view of voluntary action from being a specific brain process to being a product of stochastic activity, at least as regards the timing of action. Importantly, it is a competence model rather than a performance model: The model shows how actions and RPs could be generated without demonstrating that they actually are generated in this way in the human brain. However, the model offers a parsimonious account of action generation. Models of voluntary action as random have generally met critical reaction in philosophy as not being

part of what we mean by free will (e.g., Dennett 1984). In fact, Schurger and colleagues (2012), interestingly, leaves open the possibility that a decision to act might exist, but that this decision is taken after crossing the threshold for action, instead of at the start of the process. According to this theory, action decisions would essentially be decisions to suspend actions that are about to happen, rather than decisions to initiate them. This possibility, which is not the central part of Schurger and colleague's argument, seems more plausible for Type II than for Type I (Libet et al. 1983) actions.

Khalighinejad et al. (2018) recently developed a novel paradigm for eliciting internally generated movements at a time decided by the participant themselves [as in the work of Libet et al. (1983)], but in the context of a specific reason to act. Participants made volitional actions as a means of avoiding random and potentially long waits for stimulus onset in a perceptual decision task. Khalighinejad et al. found that these volitional skip responses were reliably preceded by a clear RP, while responses in control trials involving an external instruction to move given at the same time point after trial onset were not. The reason to act would presumably be the time saving generated by not waiting for the perceptual decision task to begin. More strikingly, they found that the variability across trials of the RP showed a marked additional reduction in the 1.5 s prior to movement onset for volitional skips, compared to the instructed control trials.

The convergence of EEG into an RP-like pattern seems, at first sight, incompatible with stochastic theories of action. However, some part of the convergence effect is clearly an artifact of aligning the data on the moment of action: The variability of a stochastic model at the moment of crossing the threshold that triggers action is defined to be zero. The fact that variability begins to decrease as early as 1.5 s prior to action seems more revealing, since it suggests that any random processes underlying action generation must be extremely slow. In fact, Khalighinejad et al. (2018) were able to model both the RP and its progressively decreasing variability using a stochastic accumulator model, but only by adding an additional parameter. This parameter involved an additional reduction in the amplitude of neural noise in the voluntary skip condition, compared to the control condition of instructed responding. This finding suggests that the stable EEG patterns associated with human volition may be due to an active process of neural noise control. The cognitive processes of attention to task and cognitive effort have likewise been identified with active processes of neural noise control (Muhammed et al. 2015).

6. THE SUBJECTIVE EXPERIENCE OF VOLITION

Table 1 highlights the fact that volition involves a characteristic subjective experience of intending and controlling actions, as well as engagement of specific neurocognitive mechanisms. Volition has a self-referential quality (Searle 1983): For an action to be volitional, one must be aware of the action and also aware that it is related to one's own processes of intending, deciding, and initiating.

Defining the content of this awareness has proved problematic, as the above quote from Wittgenstein (1953) suggests. The qualia of volition have been described as thin and elusive (Metzinger 2006), and attempts to investigate them run into the hard problem of consciousness (Chalmers 1995).

The most productive approaches to the phenomenology of volition have come from dissociations. These have shown that the phenomenology of volition can be doubly dissociated from motor activity. First, as described above, direct stimulation of the cortex of awake humans can, in specific cases, produce a phenomenal experience resembling volition, in the absence of actual movement (Desmurget et al. 2009, Fried et al. 1991). In a second dissociation, recruitment of the voluntary motor pathways can occur without phenomenology of volition. An interesting example comes from voluntary movements made under hypnotic suggestion, which are experienced



as passive. Recent studies using implicit measures rather than explicit reports (Lush et al. 2017, Oakley & Haggard 2006) support the view that suggestion reduces primary experiences of voluntariness and does not simply involve task demands favoring reports of involuntariness. A further example comes from the involuntary movements (tics) that characterize Tourette syndrome. These movements are known to emerge from activation of the cortico–striato–thalamo–cortical pathways underlying voluntary movement, yet are experienced as involuntary (Bohlhalter et al. 2006). Some reports suggest that tics are preceded by an RP that corresponds to an experience of urge to tic (Duggal & Nizamie 2002), although this remains controversial (Karp et al. 1996), and contrary findings have also been reported (Obeso et al. 1981). Taken as a whole, these examples suggest that the subjective experience of volition, although normally related to action, is doubly dissociable from action. Classical neuropsychology would conclude that the subjective experience of volition involves a different brain mechanism than does action generation. However, this logic is based on a neuropsychology of separable modules, which does not fit well with the known recurrent loop organization of the brain circuits for action generation. One might say, instead, that the experience of volition arises within the action generation loop, but from a center that can also activate independently of the loop as a whole.

The tradition of mental chronometry focuses on the subjective time of experiences, which can provide indirect information about the content of experience. The well-known experiment by Libet and colleagues (1983) has dominated discussions of volition in neuroscience and philosophy for decades. The study and its implications have been reviewed many times elsewhere (Frith & Haggard 2018, Lavazza 2016). Most previous discussions have focused on the philosophical debate between compatibilism and incompatibilism, asking whether Libet and colleagues' data invalidates the concept of conscious free will. In this section, I focus on what mental chronometry data can tell us about the subjective experience of volition and its relation to the processes and features described in **Table 1**.

Briefly, in the experiment by Libet and colleagues (1983), participants were invited to make a voluntary movement at a time of their own choice. They noted the position of a rotating spot at the time that they first felt the urge to move. The essential finding was that the conscious experience of will, which Libet and colleagues termed W judgement, occurred an average of 206 ms before movement onset. Importantly, this was markedly later than the start of the RP, which Libet and colleagues took as a marker of a specific brain process driving action initiation (for a recent alternative view, see Schurger et al. 2012). Libet and colleagues concluded that unconscious neural activity must be the cause of both our actions and our conscious experience of willing them. This result rules out Cartesian dualist views in which a conscious intention might trigger the brain activity leading to movement. However, although dualist views are widely embedded in folk psychology and general culture, they are already marginalized in modern scientific psychology. In fact, scientific psychology has long held that conscious experience is a consequence of brain activity, rather than a cause. Thus, the real impact of Libet and colleagues' work on modern psychology has not come from its rejection of dualism, but from an eliminativist reading of this work developed by Dennett & Kinsbourne (1992) and others (Banks & Isham 2009, Wegner 2002).

This reading goes beyond Libet and colleagues' (1983) assertion that the conscious experience of volition cannot trigger the initiation of action. It further suggests that conscious volition is an afterthought and not any part of the action-generating process at all. Consciousness is removed from action generation in two ways. First, the conscious experience is generated after action has been initiated, as the inferred cause of one's own actions (Banks & Isham 2009, Wegner 2002). Second, volition is seen as a post hoc interpretation of events, which structures the conscious experience of action, rather than as a real-time perception or readout of a process of action generation.

Conscious volition would become an interesting cognitive artifact in the brain's attempt to provide a coherent narrative of events, rather than a bona fide mental state or process. There is no doubt that such narrative restructuring of experience exists and plays a powerful role in experience of action. However, to claim that all experience of volition is merely post hoc narrative, as some psychologists apparently do (Banks & Isham 2009, Wegner 2002), may be too strong—people at least sometimes experience volition by the same perception-like mechanisms that allow them to perceive other sensory and motor signals, such as sensations. In the case of volition, the source of the signal is central rather than peripheral, but the experiences have similar structures and potentially even similar functions. The neuropsychological cases described above offer some evidence against strong reconstructionism. A strong reconstructionist would predict that a person with Tourette syndrome should feel their tics as voluntary (they do not) and would predict that experience of volition without action should not occur (yet it can, at least with artificial stimulation; see Fried et al. 1991).

Some recent psychophysical studies confirm an experience of volition independent of post hoc narrative. Matsushashi & Hallett (2008) asked participants to make occasional movements; however, if (and only if) they heard a randomly presented tone occurring after the time at which they first developed the intention to move, then they were asked to desist from moving for a few seconds. By analyzing the distribution of action times relative to tone time, they could estimate how long before a movement participants could detect the intention, if probed by the tone. Their estimate of -1.42 s was substantially earlier than Libet and colleagues' (1983) -206 ms. Matsushashi & Hallett suggested two levels of conscious experience of volition. At the lower level, participants are not yet spontaneously aware that they intend to act. Nevertheless, they are latently aware, because they will acknowledge the intention if specifically probed. At a higher level, participants spontaneously become aware of the intention to act. If the experience of volition were merely a reconstruction based on the moment of one's action, then the substantial time difference between these two measures would be difficult to explain. In fact, reconstructionist accounts cannot readily explain how participants could make systematic responses to the probe tone based on their own experience of volition.

In summary, in this section, I argue that a bona fide subjective experience of volition may exist. In addition to this direct experience, the brain's general reinterpretative mechanisms construct a coherent flow of one's own volitional actions. Reconstruction doubtless occurs and plays a crucial role in conscious experience generally. However, reconstructionist views have led to eliminativist readings of the Libet et al. (1983) experiment, which are oversold. Convincing neurocognitive evidence suggests that volition has a genuine preinterpretational experiential aspect, and that this experience is a product of premovement cortical activity.

7. TELEOLOGY IN HUMAN VOLITION

In this section, I discuss the aspect of volitional actions that aims at goals. Typically, humans and other animals generate volitional actions for a reason, namely to bring about some desirable outcome or goal state. The goal becomes the purpose (in Greek, *telos*) of the action. Sometimes the goal is remote in space and time when the action is initiated. In that case, the goal-directed action is internally generated, based on an internal representation of the goal state, rather than being sensorially guided as a response to a goal stimulus.

Relatively few studies on volitional action generation have provided a clear reason or goal for action. Studies in the tradition of Libet and colleagues (1983) have generally used the instruction, "Move when you feel the urge to do so," while those in the tradition of Kornhuber & Deecke (1965) have often used a self-paced instruction, "Make an action approximately every N seconds."



Both paradigms seem thin and artificial compared with the range and complexity of action choices that people confront in daily life. Another tradition, originating in reward-based decision making, offers clear choices between alternative actions that typically differ in value or affect (for a review, see Hunt & Hayden 2017). However, when the values of actions are known, the representation of the optimal action may appear almost like an instruction, and the internal generation feature of volitional action becomes elusive. In some views, all affective guidance of behavior reduces to primitive approach and avoid behaviors, triggered by appetitive and aversive stimuli, respectively (Gray & McNaughton 2003). It can be difficult to distinguish between actions that are guided by internal representations of valenced outcomes, which might count as volitional, and those that are guided by more or less explicit representations of valenced stimuli, which might seem less volitional. Moreover, the reward-guided decision making literature has a tendency to view actions merely as neutral reports of a hidden valence variable, rather than as comprising a distinct neurocognitive domain in its own right (for a notable exception, see Guitart-Masip et al. 2011).

The characteristic marker of goal-directed action is that the action is merely a means toward achieving an outcome, the goal. The goal is more important than, and hierarchically superior to, the motor action used to achieve it (Pacherie 2008, Thorndike 1927). The notion of ideomotor action (James 1890, Prinz 1987) has also been used to capture this predominance of outcomes over actions. An elegant insight into this hierarchical relationship comes from Hebb's (1949) concept of motor equivalence. Hebb argued that there are typically several possible movements that will result in a given goal or outcome, and that these movements are, in essence, all equivalent, to the extent that they achieve the same progress toward the goal. The animal cares about what happens, rather than about how it makes something happen.

To produce goal-directed action, the brain needs three essential processes: an action selection mechanism to select the action likely to produce the goal, a motor output mechanism to execute the action, and a monitoring mechanism to confirm that the goal has been achieved. The last process is particularly important if achieving the goal triggers some further action, as in a complex action sequence (Shallice 1988), or if it is useful to update the action–outcome link for potential future use. Importantly, linking these three mechanisms together allows healthy adult humans, and to a lesser extent other animals, to radically transform the environment in which they live. A rat can press a lever to receive food, but humans can build houses to have shelter or sow seeds to harvest food. The human brain's capacity for goal-directed action has profoundly transformed the environment, to the extent that, as I look around me now, everything that I see is a product of human agency, rather than a natural object.

I note above that subjective experience is a core feature of human volition. This applies equally to the teleology aspect of volitional action. The sense of agency refers to the subjective experience that one controls one's own actions and, through them, causes outcomes in the external world. This experience is so ubiquitous that it can be difficult to isolate for experimental study. Moreover, asking people directly whether they are responsible for an outcome is notoriously problematic, because of self-serving bias (Bandura 1982): People routinely report agency over positive events but deny it over negative events, regardless of whether they actually caused those events. However, implicit measures of sense of agency over outcomes have been developed based on time perception. In the intentional binding effect, the perceived interval between a volitional key-press action and an outcome such as a brief auditory tone is misperceived. The action is perceptually shifted toward the outcome, and the outcome is perceptually shifted toward the action that caused it. This amounts to an enhanced perceptual association between action and outcome, which presumably highlights the experience that one's action causes the outcome [Hume 1978 (1888)]. Crucially, this binding effect is found only for outcomes that follow one's own volitional actions and is absent or reversed when the voluntary key press is replaced by a passive movement (Haggard et al. 2002).

The intentional binding effect has previously been reviewed elsewhere as a potential measure of normal and abnormal sense of agency (Haggard 2017). Therefore, I focus in this section on recent evidence that this effect indeed tracks the operation of the three critical mechanisms for volitional control of goal-directed action described above: selecting between alternative actions, causing affectively relevant changes in the external environment, and monitoring those changes. These findings broadly confirm that intentional binding is a valid marker of sense of agency during goal-directed action.

First, several studies confirm that the ability to choose among several possible actions increases intentional binding. This suggests that action selection indeed makes an important contribution to sense of agency. Interestingly, this effect of free choice was found both when all possible actions led to the same outcome (motor equivalence; Barlas & Obhi 2013) and when they led to different outcomes (Barlas et al. 2018). Another study compared conditions in which the same action choices either did or did not predict action outcomes (Beck et al. 2017). Intentional binding was stronger in the former than in the latter case. Caspar and colleagues (2016) reported an analogous difference between binding effects when deciding for oneself and when coercively instructed by someone else.

Second, several studies have reported stronger intentional binding for positive outcomes than for negative outcomes. This result was found for emotional vocalizations (Yoshie & Haggard 2013), but only when the valence was a predictable consequence of the action (Yoshie & Haggard 2017). Stronger intentional binding was also found for less versus more painful laser stimuli triggered by voluntary actions (Beck et al. 2017) and for positive versus negative financial outcomes (Takahata et al. 2012). Interestingly, no difference in intentional binding was found when the negative versus neutral outcomes of one's action were experienced by a third party, rather than by oneself (Caspar et al. 2016).

Reinforcement guides our choices regarding what action to make. Therefore, Di Costa et al. (2017) recently embedded measures of intentional binding within a probabilistic reversal learning task to investigate how the sense of agency for volitional actions varied with reward-guided learning. Interestingly, Di Costa and colleagues did not find reliable differences in intentional binding between rewarded and unrewarded actions. However, they found a small but reliable effect related to error-based learning. The perceived time of actions was more markedly shifted toward the ensuing outcome on trials following a previous action error than on trials following a previous action that evoked a reward. This posterror agency boost was found only when action–outcome relations were stable, i.e., when people could learn which action to select. This result has important implications for the teleology of volition. The sense of agency depends on the motivation to try to make a difference in the external world through one's actions, as well as on whether one succeeds in making the intended difference or not. Trying, or conation, has lost the central place that it once had in experimental psychology (Haggard 2001), perhaps because of an understandable preference for clear instructions and defined experimental paradigms. I would like to emphasize the adaptive value of combining outcome-directed volitional action with reinforcement. This combination allows humans and other animals to generate innovative actions, to experience how their own actions can transform their environment, and to favor those actions that produce positive outcomes. Relatively simple cognitive mechanisms of this kind may be sufficient to explain the wide range and complex organization of human action, up to and including the socially oriented actions that we collectively describe as civilization (Clark 1969).

8. WIDER SIGNIFICANCE: VOLITION AND RESPONSIBILITY

Most human societies have the concept that individuals are responsible to others for the effects of their actions. This notion captures the obvious fact that one's own actions may have



important impacts on others, and that these impacts should guide individual volition. The golden rule of ethics, for example, can be seen as a constraint on individual volition: Act only according to that maxim whereby you can, at the same time, will that it should become a universal law [Kant 1993 (1785)]. Without such constraints, each individual might use their capacity for volitional action in a way that potentially harms others and themselves: Unbridled volition may be dangerous. To enforce this constraint, societies typically punish individuals who commit wrong actions and may also reward individuals who do the right thing, often through a normative system of laws. For such processes of social guidance or social constraint to work, individuals must understand that they take actions, that these actions may have important outcomes, and that the likely outcomes of their actions must drive their action choices (Law Comm. 2013). That is, they must have a neurocognitive capacity to grasp the consequences of their actions, and they must deploy this capacity at the moment of controlling the action. These conditions correspond to the teleology of volition and the generativity of volition, respectively. For example, I am responsible for the immediate and foreseeable consequences of my actions, but I am not responsible, or less responsible, for outcomes of my action that I could not reasonably foresee. Likewise, I am responsible for actions that I make with full awareness and prior planning, but not responsible, or less responsible, for movements that are involuntary or of which I am not aware, such as reflexive movements made in self-defense or movements made while sleepwalking (Rolnick & Parvizi 2011). In systems of law that descend from Roman law, the combined teleological and generative notions of volition are jointly captured by the concept of *mens rea*, which is a necessary condition for a criminal act. Correspondingly, the law allows two distinct defenses implying diminished responsibility. First, if one does not or cannot know the consequence of one's actions, one may be less responsible—the so-called cognitive defense. Second, if one was unaware that one was acting, one may be less responsible—the so-called volitional defense (Lacey 2010).

Volition is therefore a neurocognitive and psychic process of high societal importance. Societies work on the assumption that people are generally aware of their actions and of those actions' foreseeable consequences, since only then can they be responsible for an action that is wrong. Modern societies value volition because of its capacity for innovation but must nevertheless constrain it to ensure it serves a social purpose, rather than narrow individual gratification (Popper 1962). It may seem surprising that a neurocognitive process as high level and labile as volition could be the fundamental enabler of advanced societies. There are three possible reactions to this paradox. The first reaction points to the highly plastic nature of volition. Volition is strongly susceptible to modulation through learning (Di Costa et al. 2017), including social learning. Societies can and do readily guide individuals' volition through cultural reinforcement. Volition may be a cultural gadget par excellence (Heyes 2018). The second reaction minimizes the role of volition in society. In this view, most human actions are conditioned and habitual, rather than volitional. Social responsibility becomes a mechanism for conditioning actions to prevent undesirable spontaneous variants in human behavior. This view seems incompatible with modern ideas of individual freedom (Diderot & D'Alembert 1751, Skinner 1972), although it may have some application in the smaller groups found in some human societies (Dunbar 2018, Lévi-Strauss 1961). The third reaction denies the premise that so-called advanced societies have developed appropriate and proportionate constraints on human volition. In this view, burgeoning prison populations and widespread white-collar crime, among other social phenomena, would demonstrate the difficulty of reconciling the individual capacity for volition with social benefit. Scientific understanding of the neurocognitive processes underlying human volition seems inextricably linked to the wider societal aim of understanding and promoting human flourishing.

17.16 Haggard

Review in Advance first posted
on August 20, 2018. (Changes may
still occur before final publication.)



9. CONCLUSION

This review attempts to provide a working definition of volition and to show how neurocognitive mechanisms of the human brain may explain the generativity, subjectivity, and teleology of human voluntary action. The notion of conscious free will is deeply embedded in many human cultures and forms a scaffolding concept for modern society. This review does not address the question of whether people indeed have conscious free will: That is probably not even the important question to ask. People do, however, have some capacity for endogenous action, so eliminativist accounts of volition have perhaps been misguided and premature. The capacity for endogenous action may be limited, but it seems very important, particularly in social contexts. Scientific enquiry into human volition may be difficult, but it is also possible, and it is certainly valuable.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Preparation of this article was supported by ERC Advanced grant HUMVOL (agreement 323943), by a project grant from Leverhulme Trust (RPG-2016–378), and by a Chaire Blaise Pascal visiting professorship of the Région Île-de-France. I am grateful to Chris Frith, Shamil Chandraia, and Valentina Peña-Vivas for comments and assistance.

LITERATURE CITED

- Ach NK. 1905. *Über die Willenstätigkeit und das Denken*. Göttingen, Germ.: Vandenhoeck & Ruprecht
- Alexander GE, DeLong MR, Strick PL. 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9:357–81
- Bandura A. 1982. Self-efficacy mechanism in human agency. *Am. Psychol.* 37(2):122–47
- Banks WP, Isham EA. 2009. We infer rather than perceive the moment we decided to act. *Psychol. Sci.* 20(1):17–21
- Barandiaran XE, Paolo E, Di Rohde M. 2009. Defining agency: individuality, normativity, asymmetry, and spatio-temporality in action. *Adapt. Behav.* 17(5):367–86
- Barlas Z, Hockley WE, Obhi SS. 2018. Effects of free choice and outcome valence on the sense of agency: evidence from measures of intentional binding and feelings of control. *Exp. Brain Res.* 236(1):129–39
- Barlas Z, Obhi SS. 2013. Freedom, choice, and the sense of agency. *Front. Hum. Neurosci.* 7:514
- Beck B, Di Costa S, Haggard P. 2017. Having control over the external world increases the implicit sense of agency. *Cognition* 162:54–60
- Boccardi E, Della Sala S, Motto C, Spinnler H. 2002. Utilisation behaviour consequent to bilateral SMA softening. *Cortex* 38(3):289–308
- Bohlhalter S, Goldfine A, Matteson S, Garraux G, Hanakawa T, et al. 2006. Neural correlates of tic generation in Tourette syndrome: an event-related functional MRI study. *Brain* 129(8):2029–37
- Brass M, Haggard P. 2008. The what, when, whether model of intentional action. *Neuroscientist* 14(4):319–25
- Caspar EA, Christensen JF, Cleeremans A, Haggard P. 2016. Coercion changes the sense of agency in the human brain. *Curr. Biol.* 26:585–92
- Chalmers DJ. 1995. Facing up to the problem of consciousness. *J. Conscious. Stud.* 2(3):200–19
- Clark K. 1969. *Civilisation: A Personal View*. New York: Harper & Row
- Davidson D. 1963. Actions, reasons, and causes. *J. Philos.* 60(23):685–700
- Della Sala S, Marchetti C, Spinnler H. 1991. Right-sided anarchic (alien) hand: a longitudinal study. *Neuropsychologia* 29(11):1113–27



- Deiber MP, et al. 1991. Cortical areas and the selection of movement: a study with positron emission tomography. *Exp. Brain Res.* 84(2):393–402
- Dennett DC. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press
- Dennett DC, Kinsbourne M. 1992. Time and the observer: the where and when of consciousness in the brain. *Behav. Brain Sci.* 15(2):183–201
- Desmurget M, Reilly KT, Richard N, Szathmari A, Mottolese C, Sirigu A. 2009. Movement intention after parietal cortex stimulation in humans. *Science* 324(5928):811–13
- Di Costa S, Théro H, Chambon V, Haggard P. 2017. Try and try again: post-error boost of an implicit measure of agency. *Q. J. Exp. Psychol.* 71:1584–95
- Diderot D, D'Alembert JIR. 1751. *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*. Paris: Briasson
- Dolan RJ, Dayan P. 2013. Goals and habits in the brain. *Neuron* 80(2):312–25
- Duggal HS, Nizamie SH. 2002. Bereitschaftspotential in tic disorders: a preliminary observation. *Neurol. India* 50(4):487–89
- Dunbar RIM. 2018. The anatomy of friendship. *Trends Cogn. Sci.* 22(1):32–51
- Duverne S, Koehlin E. 2017. Rewards and cognitive control in the human prefrontal cortex. *Cereb. Cortex* 27(10):5024–39
- Filevich E, Kühn S, Haggard P. 2012. Intentional inhibition in human action: the power of “no.” *Neurosci. Biobehav. Rev.* 36(4):1107–18
- Fried I, Katz A, McCarthy G, Sass KJ, Williamson P, et al. 1991. Functional organization of human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.* 11(11):3656–66
- Fried I, Mukamel R, Kreiman G. 2011. Internally generated preactivation of single neurons in human medial frontal cortex predicts volition. *Neuron* 69(3):548–62
- Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. 2015. Active inference and epistemic value. *Cogn. Neurosci.* 6(4):187–214
- Frith CD. 2014. *The Cognitive Neuropsychology of Schizophrenia*. New York: Psychol. Press
- Frith CD, Friston K, Liddle PF, Frackowiak RS. 1991. Willed action and the prefrontal cortex in man: a study with PET. *Proc. R. Soc. Lond. B* 244(1311):241–46
- Frith CD, Haggard P. 2018. Volition and the brain: revisiting a classic experimental study. *Trends Neurosci.* In press
- Ghosh A, Rothwell J, Haggard P. 2014. Using voluntary motor commands to inhibit involuntary arm movements. *Proc. R. Soc. B* 281(1794):20141139
- Gold JIJ, Shadlen MMN. 2007. The neural basis of decision making. *Annu. Rev. Neurosci.* 30:535–74
- Gray JA, McNaughton N. 2003. *The Neuropsychology of Anxiety: An Enquiry Into the Function of the Septo-Hippocampal System*. Oxford, UK: Oxford Univ. Press. 2nd ed.
- Guitart-Masip M, Fuentimilla L, Bach DR, Huys QJ, Dayan P, et al. 2011. Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J. Neurosci.* 31(21):7867–75
- Haggard P. 2001. The psychology of action. *Br. J. Psychol.* 92(1):113–28
- Haggard P. 2017. Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18(4):196–207
- Haggard P, Clark S, Kalogeras J. 2002. Voluntary action and conscious awareness. *Nat. Neurosci.* 5(4):382–85
- Haggard P, Eimer M. 1999. On the relation between brain potentials and the awareness of voluntary movements. *Exp. Brain Res.* 126(1):128–33
- Halsband U, Matsuzaka Y, Tanji J. 1994. Neuronal activity in the primate supplementary, pre-supplementary and premotor cortex during externally and internally instructed sequential movements. *Neurosci. Res.* 20(2):149–55
- Hebb D. 1949. *The Organization of Behavior: A Neuropsychological Theory*. Hoboken, NJ: Wiley
- Heyes C. 2018. *Cognitive Gadgets: The Cultural Evolution of Thinking*. Cambridge, MA: Harvard Univ. Press
- Hume D. 1978 (1888). *A Treatise of Human Nature*, ed. LA Selby-Bigge, PH Nidditch. Oxford: Oxford Univ. Press. 2nd ed.
- Hunt LT, Hayden BY. 2017. A distributed, hierarchical and recurrent framework for reward-based choice. *Nat. Rev. Neurosci.* 18(3):172–82

- Jahanshahi M, Jenkins IH, Brown RG, Marsden CD, Passingham RE, Brooks DJ. 1995. Self-initiated versus externally triggered movements. I. An investigation using measurement of regional cerebral blood flow with PET and movement-related potentials in normal and Parkinson's disease subjects. *Brain* 118(4):913–33
- James W. 1890. The principles of psychology. *J. Hist. Philos.* 21(2):270–72
- Johansson RS, Lemon RN, Westling G. 1994. Time-varying enhancement of human cortical excitability mediated by cutaneous inputs during precision grip. *J. Physiol.* 481(3):761–75
- Kant I. 1993 (1785). *Grounding for the Metaphysics of Morals*. Indianapolis, IN: Hackett Publ. 3rd ed.
- Karp BI, Porter S, Toro C, Hallett M. 1996. Simple motor tics may be preceded by a premotor potential. *J. Neurol. Neurosurg. Psychiatry* 61(1):103–6
- Khalighinejad N, Schurger A, Desantis A, Zmigrod L, Haggard P. 2018. Precursor processes of human self-initiated action. *NeuroImage* 165:35–47
- Kornhuber HH, Deecke L. 1965. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Arch. Eur. J. Physiol.* 284(1):1–17
- Kritikos A, Breen N, Mattingley JB. 2005. Anarchic hand syndrome: bimanual coordination and sensitivity to irrelevant information in unimanual reaches. *Cogn. Brain Res.* 24(3):634–47
- Lacey N. 2010. Psychologising Jekyll, demonising Hyde: the strange case of criminal responsibility. *Crim. Law Philos.* 4(2):109–33
- Lavazza A. 2016. Free will and neuroscience: from explaining freedom away to new ways of operationalizing and measuring it. *Front. Hum. Neurosci.* 10:262
- Law Comm. 2013. *Criminal liability: insanity and automatism*. Rep., Law Comm., London. http://www.lawcom.gov.uk/app/uploads/2015/06/insanity_discussion.pdf
- Lévi-Strauss C. 1961. *Tristes tropiques*. New York: Criterion Books
- Levy N. 2011. Expressing who we are: moral responsibility and awareness of our reasons for action. *Anal. Philos.* 52(4):243–61
- Libet B, Gleason CA, Wright EW, Pearl DK. 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(3):623–42
- Lush P, Caspar EA, Cleeremans A, Haggard P, Magalhães De Saldanha da Gama PA, Dienes Z. 2017. The power of suggestion: posthypnotically induced changes in the temporal binding of intentional action outcomes. *Psychol. Sci.* 28(5):661–69
- Matsushashi M, Hallett M. 2008. The timing of the conscious intention to move. *Eur. J. Neurosci.* 28(11):2344–51
- Metzinger T. 2006. Conscious volition and mental representation: toward a more fine-grained analysis. In *Disorders of Volition*, ed. N Sebanz, W Prinz, pp. 19–48. Cambridge, MA: MIT Press
- Mischel W, Ebbesen EB, Zeiss AR. 1972. Cognitive and attentional mechanisms in delay of gratification. *J. Personal. Soc. Psychol.* 21(2):204–18
- Misirlişoy E, Haggard P. 2014. Veto and vacillation: a neural precursor of the decision to withhold action. *J. Cogn. Neurosci.* 26(2):296–304
- Muhammed K, Manohar S, Husain M. 2015. Mechanisms underlying apathy in Parkinson's disease. *Lancet* 385:S71
- Oakley DA, Haggard P. 2006. The timing of brain events: authors' response to Libet's "Reply." *Conscious. Cogn.* 15(3):548–50
- Obeso JA, Rothwell JC, Marsden CD. 1981. Simple tics in Gilles de la Tourette's syndrome are not prefaced by a normal premovement EEG potential. *J. Neurol. Neurosurg. Psychiatry* 44(8):735–38
- Pacherie E. 2008. The phenomenology of action: a conceptual framework. *Cognition* 107(1):179–217
- Parvizi J, Rangarajan V, Shirer WR, Desai N, Greicius MD. 2013. The will to persevere induced by electrical stimulation of the human cingulate gyrus. *Neuron* 80(6):1359–67
- Passingham RE. 1987. Two cortical systems for directing movement. *Ciba Found. Symp.* 132:151–64
- Passingham RE. 1993. *The Frontal Lobes and Voluntary Action*. Oxford, UK: Oxford Univ. Press
- Passingham RE, Bengtsson SL, Lau HC. 2010. Medial frontal cortex: from self-generated action to reflection on one's own performance. *Trends Cogn. Sci.* 14(1):16–21



- Picard N, Strick PL. 1996. Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* 6(3):342–53
- Popper K. 1962. *The Open Society and Its Enemies*, Volume I: *The Spell of Plato*. New York: Routledge
- Prinz W. 1987. Ideo-motor action. In *Perspectives on Perception and Action*, ed. H Heuer, AF Sanders, pp. 47–76. Hillsdale, NJ: Lawrence Erlbaum Assoc.
- Rohrbaugh JW, Gaillard AWK. 1983. Sensory and motor aspects of the contingent negative variation. *Adv. Psychol.* 10:269–310
- Rolnick J, Parvizi J. 2011. Automatisms: bridging clinical neurology with criminal law. *Epilepsy Behav.* 20(3):423–27
- Schurger A. 2018. Specific relationship between the shape of the readiness potential, subjective decision time, and waiting time predicted by an accumulator model with temporally autocorrelated input noise. *eNeuro* 5(1):ENEURO.0302-17.2018
- Schurger A, Sitt JD, Dehaene S. 2012. An accumulator model for spontaneous neural activity prior to self-initiated movement. *PNAS* 109(42):E2904–13
- Schüür F, Haggard P. 2011. What are self-generated actions? *Conscious. Cogn.* 20(4):1697–704
- Searle JR. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge, UK: Cambridge Univ. Press
- Shallice T. 1988. *From Neuropsychology to Mental Structure*. Cambridge, UK: Cambridge Univ. Press
- Sherrington C. 1906. *The Integrative Action of the Nervous System*. New Haven, CT: Yale Univ. Press
- Sirigu A, Daprati E, Ciancia S, Giraux P, Nighoghossian N, et al. 2004. Altered awareness of voluntary action after damage to the parietal cortex. *Nat. Neurosci.* 7(1):80–84
- Skinner BF. 1972. *Beyond Freedom and Dignity*. New York: Bantam Books
- Sumner P, Nachev P, Morris P, Peters AM, Jackson SR, et al. 2007. Human medial frontal cortex mediates unconscious inhibition of voluntary action. *Neuron* 54(5):697–711
- Takahata K, Takahashi H, Maeda T, Umeda S, Suhara T, et al. 2012. It's not my fault: postdictive modulation of intentional binding by monetary gains and losses. *PLOS ONE* 7(12):e53421
- Thorndike EL. 1927. The law of effect. *Am. J. Psychol.* 39(1/4):212–22
- Verleger R, Haake M, Baur A, Śmigajewicz K. 2016. Time to move again: Does the Bereitschaftspotential covary with demands on internal timing? *Front. Hum. Neurosci.* 10:642
- Voss M, Chambon V, Wenke D, Kühn S, Haggard P. 2017. In and out of control: brain mechanisms linking fluency of action selection to self-agency in patients with schizophrenia. *Brain* 152(8):33–40
- Wegner DM. 2002. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press
- Wittgenstein L. 1953. *Philosophical Investigations*. Hoboken, NJ: Wiley. 3rd ed.
- Yoshie M, Haggard P. 2013. Negative emotional outcomes attenuate sense of agency over voluntary actions. *Curr. Biol.* 23(20):2028–32
- Yoshie M, Haggard P. 2017. Effects of emotional valence on sense of agency require a predictive model. *Sci. Rep.* 7:8733
- Zapparoli L, Seghezzi S, Scifo P, Zerbi A, Banfi G, et al. 2018. Dissecting the neurofunctional bases of intentional action. *PNAS* 115(28):7440–45