Voluntary action requires consciousness

Tyler J. Adkins

University of Michigan

Cognitive Science

Honors Thesis

Voluntary action requires consciousness

Consider the following cases:

1. On the morning of January 1 1961, Willis Boshears strangled a young woman to death. It was determined beyond reasonable doubt that he performed this violent act, but he was nevertheless acquitted of the charge of murder (Shearer, 1961).
2. On the morning of May 24th, 1987, Kenneth James Park drove over 20 km to his mother-in-law's home and stabbed her to death. Once again, it was determined beyond reasonable doubt that he performed this violent act, but he too was acquitted on the charge of murder (Broughton, 1994).
3. On a July night in 2008, Brian Thomas strangled his wife to death. And again, although there was no doubt that he strangled his wife, a jury found him to be not guilty of murder (Morris, 2009).

In cases 1, 2, and 3, the defendant performed a morally repugnant, violent action against another person, yet the courts determined that the defendant was not guilty of a crime. In all three cases, the expressed reason why the defendant was not guilty was that the defendant did not *voluntarily* perform the violent action because the defendant was not consciously aware of any aspect of his action. A key part of this line of thought is captured by the following principle:

V: Voluntary action requires consciousness.

The following schema represents the gist of my argument for **V**, where ➔ is transitive and means "requires":

Voluntary action ➔ Voluntary attention ➔ Consciousness

I will devote section one of this essay to a discussion of attention. In section two, I will provide support for the hypothesis that voluntary attention is necessary for voluntary action. In section three, I will discuss consciousness. In section four, I will provide support for the hypothesis that consciousness is necessary for voluntary attention. In section five, I will use the key takeaways from sections one through four to argue that voluntary action requires consciousness. In section six, I will discuss some research on the subjective feeling of voluntariness and the causal connection between consciousness and action. The purpose of this discussion is to show that the feeling of voluntariness may be the phenomenal counterpart to the functional role that consciousness plays in voluntary action. In section seven, I will present a theoretical narrative about the cognitive processes involved in voluntary action which involves goals, consciousness, attention, and action. In section seven, I will unpack some experimental predictions and philosophical features, as well as theoretical flaws, of this narrative about voluntary action.

§1: Attention

The amount of information available for processing by the human brain at a given time is astounding. For example, it has been estimated that the information transfer rate of the human eyes is 10^8-10^9 bits per second (Borji & Itti, 2013; K. Koch et al., 2006)! But the human brain does not simply process visual information. It is also bombarded with information feeds from transducers in the other sensory organs, such as feeds of auditory, tactile, olfactory, and proprioceptive signals. Furthermore, there are vast numbers of endogenous signals, originating from within the brain, which represent thoughts, emotions, desires, intentions, and other mental states. Even if it were biologically possible for the brain to deeply process all this information—such that the information has a significant influence on subsequent behavior and mentation—doing so would yield highly incoherent memories, experiences, and behaviors. So, to respond coherently and intelligently to this mass amount of information, the brain must *select* a proper subset of the available information for enhanced processing. A basic conception of attention, one that is shared by most contemporary theories of the process, identifies it with this process of selection (for a review of such theories, see Carrasco, 2011).

One neurally plausible theory of attention treats it as a process of 'biased competition' (Desimone & Duncan, 1995; Webb & Graziano, 2015). According to this account, signals *compete* with one another for enhanced processing and the signals that win this competition are *attended*. The existence of this sort of competitive processing is consistent with the organization and patterns of activity of cells in the retina (Hartline, Wagner, & Ratliff, 1956; Kuffler, 1953) and visual cortex (Kastner & Ungerleider, 2000; Moran & Desimone, 1985; Reynolds, Chelazzi, & Desimone, 1999). An important aspect of this competition for enhanced processing is that it is biased. A signal is selected for enhanced processing—that is, attended—if and only if its strength exceeds some threshold. Bottom-up features of stimuli and top-down goals both contribute to the amplification of signals beyond this threshold. In this sense, bottom-up features and top-down goals can *bias* the signal competition by *amplifying* the strength of some signals, making them more likely than other signals to be selected for enhanced processing.

Attention—that is, the selection of signals for enhanced processing—can be involuntary or voluntary. In the case of involuntary attention—sometimes referred to as *attentional capture* (Wu, 2011)—the signal competition described above is influenced by "bottom-up" properties of the stimulus, such as its brightness, contrast, or intensity. In the case of voluntary attention, the signal competition is biased by a person's goals, intentions, desires, plans, and the like. For simplicity, I will hereafter refer to all such mental states as "goals". In terms of neural correlates, studies of human and primate brains show that a bilateral network including the dorsal posterior parietal cortex (PPC) and frontal cortex is engaged during the deployment of top-down biases of attention (Corbetta & Shulman, 2002). Additionally, studies using electroencephalography (EEG) show an increase in activity in the gamma range (30-70 Hz) during voluntary, but not involuntary, attention (Landau, Esterman, Robertson, Bentin, & Prinzmetal, 2007). The existence of dissociable neural correlates of these phenomena warrants us in being more confident that they are really distinct.

§2: Voluntary action requires voluntary attention

A first-pass definition of voluntary action is:

> **VA:** A person *P* voluntarily performs an action *A* if and only if *P* performs *A* and *P* possesses the goal to perform *A*.

One problem with this definition is that it does not specify the type of causal connection that obtains between a goal and action in the case of voluntary action. Intuitively, it cannot be a mere *accident* that one's goal to perform *A* leads one to perform *A* (Wu, 2016). An example which pumps this intuition is one in which a man has a goal to murder his uncle. The man then gets very nervous because of this goal, so nervous that he accidentally runs over a pedestrian who happens to be his uncle, with his car (Chisholm, 1966). Despite his having the goal to kill his uncle, and his succeeding in doing so, it does not seem that he voluntarily killed his uncle. The explanation for this is that his goal to kill was not causally related to the killing "in the right sort of way" (Wu, 2016). So, definition **VA** is inadequate.

Wayne Wu's (2016) solution to the problem of specifying the type of causal connection between goal and action, which must obtain for an action to be voluntary, is to incorporate *attention* into the story. In the case of voluntary action, goals and actions are connected via *voluntary attention*. Voluntary attention involves the enhanced processing of goal-relevant signals—processing which enables these signals to drive behavior. So, according to Wu, the missing link in the causal chain in Chisholm's example is attention. On this view, when one performs a voluntary action *A*, one possess a goal *G* to perform *A* and *G* influences attention to enhance the processing of *G*-relevant signals, which in turns enables these signals to drive behavior. Attention is needed for the systematic coordination of goal-oriented action. That is, voluntary action requires voluntary attention.

Consistent with the above ideas is an extensive amount of literature which demonstrates the effects of attention on the control of bodily movements (Lohse, Jones, Healy, & Sherwood, 2014). For example, attentional focus has been shown to affect the activity of muscle cells measured using electromyography (EMG) (Lohse & Sherwood, 2012; Vance, Wulf, Töllner, McNevin, & Mercer, 2004; Zachry, Wulf, Mercer, & Bezodis, 2005). Additionally, attentional focus has been shown to affect energetic costs of movement as measured by oxygen consumption (Schücker, Hagemann, Strauss, & Völker, 2009), the variability of the kinematics of joints (Lohse, Sherwood, & Healy, 2010), and the control of gait (Woollacott & Shumway-Cook, 2002). Additionally, there are numerous studies showing that an external focus of attention—directed at the *intended outcome* of the movement—provides an advantage over an internal focus of attention—directed at the *body or movement*—in the learning and performance of a wide variety of motor skills (Bell & Hardy, 2009; Lohse et al., 2010; Wulf, 2007, 2013; Wulf, Mcconnel, Gärtner, Schwarz, & Gartner, 2002; Wulf & Su, 2007; Lohse & Sherwood, 2012; Zachry et al., 2005).

Lohse, Healy, and Sherwood (2014) provide a hypothesis of the role of attention in motor control based on an optimal control framework proposed by Sholz and Schöner (1999). This framework posits a space of all possible movements and divides the space four ways. The first division is between *goal-relevant dimensions*—movement dimensions (e.g., joint angles or velocities) which influence whether the movement is successful—and *redundant dimensions*—movement dimensions which do not influence whether the movement is successful. They add a further distinction between the *uncontrolled manifold*—the set of movement dimensions which can vary—and the *controlled manifold*—the set of movement dimensions which are not allowed to vary. Within this framework, implementing a strategy for successful movement amounts to aligning the controlled manifold with the set of goal-relevant dimensions and allowing the goal-relevant dimensions to vary leads to diminished performance (Lohse et al., 2014).

Using this framework, Lohse, Healy, and Sherwood (2014) propose the following hypothesis to explain the effects of attention on motor variability and performance: changing one's focus of attention helps the motor system decide which movement dimensions are goal-relevant and which are redundant, and hence which movement dimensions to control (i.e., prevent from varying). An external focus of attention leads to increased control of movement dimensions which are *in fact* goal-relevant (e.g., correlations between kinematics of distinct joints), while an internal focus of attention leads to increased control of bodily dimensions which are in fact not goal-relevant (e.g., kinematics of individual joints). This hypothesis is supported by their observation that an external focus of attention is associated with increased variability (i.e., decreased control) of individual joint angles and velocities but increased correlation among these dimensions over time, while an internal focus is associated with the opposite effects (Lohse et al., 2014).

All this suggests that attention has modulatory effects on the activity of the motor system. Goal-oriented attention—that is, voluntary attention—facilitates performance by specifying the dimensions of movement over which it is optimal to enact control as well as the dimensions in which it is optimal to allow for variability. The upshot of this is that voluntary attention is necessary for the optimal control of goal-oriented or voluntary movement.

§3: Consciousness

Ned Block famously distinguishes between two types of consciousness: access and phenomenal (Block, 1990, 1992, 1995). *Phenomenally* conscious mental states are associated with internal, subjective experiences: there is *something it is like* to be in them (Nagel, 1974). For example, consider the difference between what it is like to see red and what it is like to see blue: this difference is in terms of *phenomenal* properties. On the other hand, *accessibly* conscious mental states are characteristically available as inputs to executive functions, such as reasoning, speech, and action (Block, 1995). For example, my ability to report that I see a green Christmas tree requires that my perceptual representation of the tree is accessibly conscious.

While attempts have been made to interpret certain behavioral results as successfully disentangling phenomenal consciousness and access consciousness (e.g., Block, 2005), there is good reason to be skeptical about this possibility. The reason for skepticism is that the only way to measure phenomenal consciousness is via reports about the content of an experience. These reports might be verbal, such as saying “I see the tree” or “I am experiencing the red circle”, motoric, such as pressing a button if a stimulus is visible, or they may take the form of wagers that represent one’s confidence about the content of one’s experience, such as betting $5 that the stimulus is present (Webb & Graziano, 2015). For each variety of report, the *phenomenal*—that is, experiential—content of a mental representation can be measured if and only if it is the representation is *accessible*—that is, available for report.

Block (2005) argues that *introspective* reports should not be the “gold standard” for measuring phenomenal consciousness, since introspection is a species of metacognition and young children and non-human animals have deficient capacities for metacognition compared to adult humans. Block advocates for the use of *environmental* reports, which require a subject to report “what the subject takes to be in front of him”, as an alternative to introspective reports, which require a subject to report “what the subject is experiencing” (Block, 2005).

My response to the purported distinction between introspective and environmental reports is that “what the subject takes to be in front of him” and “what the subject is experiencing” are coextensive: they refer to the same thing. When one reports that there is no stimulus in front of him, we naturally infer from this report that he has no experience of the stimulus, and vice versa. This is because we assume the two types of reports to be describing the same thing. Therefore, if young children can provide *environmental reports,* then they possess the capacity to access the contents of their experiences (although they might be metacognitively deficient in other respects). More importantly, however, the takeaway here is that phenomenal consciousness and access consciousness cannot be experimentally disentangled. So, I will hereafter assume that the two phenomena are identical or always correlated: consciousness is always both phenomenal and accessible. On my view, each of the functional properties attributed to consciousness—such as, enabling voluntary attention or action—is associated with a distinctive subjective feeling of what it is like.

Many contemporary theorists argue that consciousness serves a sort of supervisory role, whereby it contributes to the higher-order monitoring or guidance of certain types of mental and bodily activities (Dehaene & Changeux, 2011). This view traces back at least as far as William James, who characterized consciousness as “an organ added for the sake of steering a nervous system grown too complex to regulate itself” (James, 1890). Bernard Baars, in discussions of his famous Global Workspace Theory of consciousness, frequently mentions how consciousness enables voluntary control (Baars, 1997, 2002, 2005). One result that he cites to support this claim is a study in which participants learned to control the firing patterns of a single motor neuron based on conscious auditory feedback which represents the neuron's activity (Simard & Basmajian, 1967). Baars points out that such control has never been observed with unconscious feedback (Baars, 2005).

Another interesting supervisory role for consciousness concerns its relationship with attention. Attention and consciousness almost always correlate (Mack & Rock, 1999; Prinz, 2013). Nevertheless, there is extensive evidence which demonstrates that the two processes are behaviorally and neurally dissociable (Kentridge, Nijboer, & Heywood, 2008; C. Koch & Tsuchiya, 2007; McCormick, 1997; Norman, Heywood, & Kentridge, 2013; Webb, Igelström, Schurger, & Graziano, 2016). This intimate but dissociable relationship between attention and consciousness has inspired some to theorize that consciousness is a higher-order representation or model *of* attention (Webb & Graziano, 2015). The Attention Schema Theory developed by Michael Graziano argues that when a brain attends to a stimulus and computes a higher-order representation *of* this attentional state, the brain will conclude that it is aware (Webb & Graziano, 2015). Graziano's Attention Schema Theory treats conscious awareness as a perceptual and pragmatic internal model of attention. The internal model is *perceptual* because it aims to represent attention in its current state. According to the theory, the internal model represents the *contents* of attention *as* attended. Importantly, this internal model is best understood as a *schema* because while it represents attention and its contents, it does so in a coarse-grained and approximate manner. Since the attention schema describes attention with no more detail than is necessary to facilitate the guidance or control of attention, it can be thought of as essentially *pragmatic*. In section four, I will discuss why possessing an internal model of attention would make it easier for the brain to optimally control attention.

Evidence from recent fMRI studies implicate an area of the right temporal-parietal junction (rTPJ) in the computation of an attention schema (Kelly, Webb, Meier, Arcaro, & Graziano, 2014; Webb, Igelström, et al., 2016). This brain region was consistently activated when participants were instructed to think about the attentional state of another person—that is, represent an attentional state. And most strikingly, decreasing activity in this region using Transcranial Magnetic Stimulation (TMS) leads to disruptions of conscious visual experience (operationalized as the capacity to report the presence of a white dot). These results support not only the localization of the computation of an attention schema to the rTPJ, but also the necessity of such computational processes for normal conscious visual awareness.

§4: Voluntary attention requires consciousness

A large body of behavioral evidence suggests that voluntary attention is possible only with the help of consciousness (Ivanoff & Klein, 2003; Lambert, Naikar, McLachlan, & Aitken, 1999; McCormick, 1997; Webb, Kean, & Graziano, 2016), whereas bottom-up, stimulus-driven attention dominates in the absence of consciousness (Lambert et al., 1999; McCormick, 1997; Tsushima, Sasaki, & Watanabe, 2006).

Additionally, a key prediction of the Attention Schema Theory—described in §2—is that attention should suffer severe deficits of control in the absence of consciousness. There is a handful of results consistent with this prediction, and they reveal a pattern which indicates the specific variety of attentional control that is compromised in the absence of consciousness. In a study in which participants were presented with lateralized cues which indicated a subsequent contralateral presentation of a target, the extent to which participants could resist the reflex to attend to the side of the cue was affected by whether they were consciously aware of the cue (McCormick, 1997). When participants were consciously aware of the cue, they could successfully voluntarily attend to the side where the target would be presented. However, when participants were not consciously aware of the cue, they just focused on the side where the cue was presented. In another study, it was observed that when participants perform a letter identification task with distracting peripheral stimuli, they perform *better* if they are consciously aware of the distracting stimuli than if they are not consciously aware of the distracting stimuli (Tsushima et al., 2006). In both studies, it was observed that the ability to resist bottom-up attentional capture is diminished in the absence of awareness. This suggests that the top-down, voluntary control of attention may require consciousness.

An important feature of the Attention Schema Theory is its relation to model-based dynamical systems control, an engineering framework which assumes that to achieve optimal control of a complex dynamical system, it is useful for the controller to have access to an internal model of the system being controlled (Camacho & Bordons, 2007a; Camacho & Bordons, 2007b; Webb, Kean, & Graziano, 2016). Since the Attention Schema theory treats consciousness as an internal model used for the control of attention, it predicts that the absence of consciousness will result in the deficits of attentional control analogous to the characteristic behavior of dynamical systems whose controller is missing an internal model. These deficits include a diminished capacity for task-relevant attentional focus, diminished ability to resist task-irrelevant attentional focus, and increased likelihood of attention being influenced by external factors (Graziano & Botvinick, 2002; Scheidt, Conditt, Secco, & Mussa-Ivaldi, 2005; Webb, Kean, & Graziano, 2016; Wolpert, Ghahramani, & Jordan, 1995). In a recent study, all three of these predicted attentional deficits were observed. (Webb, Kean, et al., 2016). In this study, when conscious awareness was absent, it was more difficult for participants to maintain focus on goal-relevant stimuli. Additionally, the ability to resist attentional capture by a task-irrelevant stimulus was reduced and attentional selection was more biased by the contrast of the stimulus, in the absence of consciousness. A plausible explanation for these latter findings is that the default state of attention, in the absence of top-down influence from goals and the like, is to be controlled by bottom-up biases such as contrast. If so, then it appears consciousness may be necessary for attention to be guided in a goal-oriented manner.

§5: Voluntary action requires consciousness

Based on the above discussion, I propose the following argument for the hypothesis that consciousness is necessary for voluntary action:

P1. *P* voluntarily performs *A* only if *P's* goal to perform *A* guides *A.*

P2. A goal *G* optimally guides an action only if *G* influences attention to signals associated with *G*-relevant action dimensions (*GAD)*.

P3. *G* optimally guides attention to *GAD* only if consciousness enables goals to influence attention.

C1. So, *P* voluntarily performs an action *A* only if consciousness enables goals to influence attention.

For the purposes of this paper, I assume that P1 is uncontroversial. It is difficult to imagine a scenario in which one voluntarily performs an action yet lacks any goal-like state to guide or inform this action.[1] P2 specifies the necessity of attention for voluntary action and is supported

[1] One potential class of counterexamples includes cases where an individual (i) performs some action *A1,* but (ii) never had the goal to perform *A1*, (iii) had the goal to perform a *similar but distinct* action *A2,* nevertheless (iv) seems to perform *A1* voluntarily. For example, suppose Bob intends to shoot a deer, sees a deer-looking creature and shoots it, but the creature turns out to be a deer-looking goat. In this case, Bob never had no intention to kill that goat. Yet, perhaps it is intuitive that Bob voluntarily killed the goat. My intuition is that Bob did not voluntarily kill the goat. At most, he voluntarily killed what he *knew* was there: a deer-looking creature.

by the theories and experimental results presented in section three of this essay. P3 specifies the necessity of consciousness for voluntary attention and is supported by the theories and experimental results presented in section four of this essay. C1 entails that voluntary action requires consciousness—and it is logically entailed by the conjunction of P1, P2, and P3.

§6: Feeling of voluntariness

The feeling of voluntariness refers to the subjective experience as of having *conscious control* over one's action: it is the feeling that consciousness contributes to the guidance of action (Wegner, 2003). Nearly all human actions are associated with this feeling of voluntariness and I would wager that the widespread folk-psychological belief in the existence of free will is based primarily on this feeling of voluntariness. For many decades, the common view among psychologists and neuroscientists has been that the feeling of voluntariness is illusory: that is, while it may *seem* that consciousness contributes to the guidance of action, this is often not (perhaps, *never)* the case.

One prominent hypothesis about the feeling of voluntariness, which I will call the *prospective-retroactive-agreement hypothesis* or PRA for short, is that the feeling of voluntariness arises when there is agreement between the *prospective signals* representing the goal to move and the *retroactive signals* representing the sensory consequences of the movement (Farrer et al., 2008; Haggard & Clark, 2003). Importantly, most of the retroactive signals mentioned above are *simultaneous* with the movement itself, since movements take time to complete. So, PRA is consistent with the hypothesis that the subjective feeling of voluntariness contributes to the coordination of voluntary action. PRA is primarily supported by studies which show correlations between reported feelings of *involuntariness* and *disagreement* between the intended and actual outcomes of a movement (Moore, Wegner, & Haggard, 2009; Wegner, 2003).

Numerous brain regions have been implicated in the processing of this disagreement, which is thought to underlie the feeling of involuntariness (Wegner, 2003).

Other studies on the feeling of voluntariness focus on the subjective temporal difference between the onset of a movement and the onset of its effect. One finding is that this subjective temporal difference gets smaller following a movement made “by choice”, and gets larger following a motor reflex induced by TMS (Haggard et al., 2002). The reduction of this subjective temporal difference—known as the binding effect—is often used as a proxy for measuring the feeling of voluntariness (Moore et al., 2009). In a more recent study, it was observed that when participants are not told in advance which event-onset to remember among (G) the goal, (M) the movement, or (C) the consequences of the movement, the perceived temporal difference is much smaller. In such conditions, subjects are nearly unable to distinguish between the onsets of G, M , and C (Haggard & Cole, 2007). According to Haggard and Cole, when participants are not told in advance whether they are to remember the onset of G, M, or C, participants do not direct their attention to G, M, or C. Given this, the authors infer that a lack of attention to G, M, and C leads to increased binding effects. If (a) binding effect is a valid implicit measure of feeling of voluntariness and (b) the binding effect is strongest when attentional focus is not directed to G, M, or C, then it is reasonable to infer that the most robust feelings of voluntariness arise when attention is not focused on G, M, or C. However, I think there are reasons to doubt (a) and (b).

First, there is good reason to doubt Haggard and Cole’s interpretation of their results and hence (b). It is unlikely that when the participants are not told in advance whether they should remember the onset of G, M, or C, that the participants do not attend to G, M, or C. After all, the participants still understand that they must recall the onset of one of G, M, or C, they just do not know which one. Rather than simply not attending to any of the events—a strategy which would render the task of accurately reporting the onset of one of the events difficult or impossible—it is much more likely that the participants *alternate* their attentional focus between the events. This alternating attention strategy, while not quite as poor as the strategy of *non-attention* which Haggard and Cole attribute to their participants, is likely to yield the inaccurate estimates of the onsets of G, M, and C which are observed in the study. Furthermore, if participants divide their attention between G, M, and C, this may interfere with the ordinary dynamics of attention involved in voluntary action. If so, then participants may in fact feel less conscious control—that is, a diminished feeling of voluntariness—when they are not told in advance whether they must recall the onset of G, M, or C, precisely because the normal processes of voluntary control are disrupted by the task demands.

Concerning (a), it is an open empirical question whether the binding effect is a valid measure of the feeling of voluntariness. One way to test whether the binding effect is a valid measure of the feeling of voluntariness is to check it against alternative valid measures, such as having participants judge on a scale ranging from 'no control' to 'complete control' the extent to which they thought they controlled their action (Chambon & Haggard, 2012). To my knowledge, no comparison between these two measures has been reported in print.

The most influential results in the literature on the feeling of voluntariness come from Benjamin Libet's experiments, which showed that a readiness potential in motor cortex regularly occurred 350-400 ms *before* the time at which participants became consciously aware of their goal to flex their wrist (Libet, 1985). The conclusion often drawn from this result is that the feeling of voluntariness does not initiate movement, since it always arises *after* the brain has already initiated the causal chain of neural activity that will reliably lead to action (P Haggard et al., 2002; Patrick Haggard & Eimer, 1999; Wegner, 2003). This result suggests that the feeling of voluntariness arises *prior* to the action, since participants in Libet's experiment became consciously aware of their goal to flex their wrist 100-150 ms before they flexed their wrist. But, more importantly, the Libet findings are consistent with the hypothesis that the feeling of voluntariness contributes to voluntary action by enabling goals to influence attention.[2] This contribution of consciousness could occur at any time between the onset and conclusion of an action. So, the Libet findings do not support the hypothesis that consciousness has no causal contribution to action. All results presented above are consistent with the hypothesis that the feeling of voluntariness contributes to the coordination of voluntary action.

§7: A theoretical narrative about voluntary action

What emerges from the arguments in sections five and six is a theoretical narrative about the cognitive mechanisms that jointly enable voluntary action. I will hereafter refer to this narrative as **N**. A rough draft of **N** is as follows: prior to performing a voluntary action, one has a goal about what to do. Next, this goal is used in conjunction with consciousness—understood as a useful internal model of attention—to optimally *guide* attention toward dimensions of the subsequent movement that are *goal-relevant*. Attention then enhances the processing of the signals associated with these goal-relevant movement dimensions, thereby enabling the motor system to control the limbs in a way which is optimal with respect to the initial goal. When a voluntary action is occurring, consciousness qua functional incorporates (at minimum) information about goals, the contents of attention, and the process of attention. But, perhaps at the same time, consciousness qua phenomenal has a different set of properties, such as *what it is*

[2] Another possible role for consciousness to play in voluntary action is as a mechanism by which the brain can *veto* unfolding actions before they are completed. Thanks to Eric Lormand for reminding me of this possibility.

like to perform a voluntary action. In other words, perhaps, when one is performing a voluntary action, one's consciousness qua phenomenal is simply the subjective feeling of voluntariness.

§8: Implications and shortcomings of N

S leads to some testable predictions. According to **N**, voluntary attention and consciousness are critical causal nodes that connect goals to actions. Therefore, in the absence of either of these processes, voluntary action will not be possible. Furthermore, given the way in which **N** connects voluntary action and the feeling of voluntariness, the absence of consciousness or voluntary attention should also lead to disruptions in the feeling of voluntariness. Also, to the extent that there are well-defined neural mechanisms underlying the processes of consciousness and voluntary attention, modulation of the activity of these mechanisms via TMS should affect the capacity to produce voluntary actions, as well as the feeling of voluntariness. Additionally, one should find systematically coordinated activity between regions underlying the computation of goals—such as areas of prefrontal cortex—and regions underlying the computation of a higher-order model of attention—such as the TPJ—and regions responsible for attentional control—such as areas in the dorsal PPC—during both voluntary attention and voluntary action. Further, the robustness of this coordinated neural activity should be positively correlated with the feeling of voluntariness.

N also has some interesting philosophical features. First, **N** vindicates the intuition presented at the start of this essay that voluntary action requires consciousness. Second, **N** assigns a specific causal function to consciousness, thereby rebuffing the view that consciousness is purely *epiphenomenal* (Huxley, 1874). Also, if **N** is correct, then the subjective feeling of voluntariness is normally *veridical*, rather than *illusory*: it is not merely the experience *as of* acting voluntarily, rather it is the experience *of* acting voluntarily. Another important feature of **N**

is that it provides a purely *naturalistic* conception of voluntariness, motivated by results in contemporary psychology and neuroscience. **N** also has the virtue of eschewing debates about the compatibility of voluntary action with determinism. Whether the physical processes underlying an action are deterministic and linear or stochastic and non-linear is totally irrelevant to whether the action is voluntary. If the appropriate interplay between goals, consciousness, attention, and action obtain, the action will be voluntary, regardless of whether the underlying processes are deterministic.

However, **N** is highly inadequate as a theory of voluntary action because it fails to provide specific details about the neural and computational realizations of the interactions between goals, consciousness, attention, and action. This shortcoming is partly explained by the difficulty I had in recognizing meaningful patterns of results in the contemporary literature concerning such mechanisms. For example, how is it that the brain constructs an internal model of attention? And what are the mechanisms that mediate the connection between goals and attention? In the future, it would be worthwhile to try to develop computational models which make precise predictions about statistical patterns of behavioral and neuroimaging data.

Also, there is room for skepticism about whether voluntary action according to **S** is the sort of thing that can ground moral responsibility. After all, there are cases in which an action would be declared voluntary by **S**, but which nevertheless are not actions for which an agent should be held morally responsible. The most obvious type of case concerns actions which are coerced. For example, suppose some guy Bob is a store and a fellow shopper walking near Bob grabs him, brandishes a handgun, and tells him that he must steal an item from the store, or else he will be shot. Bob, out of fear for his life, decides—that is, forms a goal—to steal the item. This goal, with the help of consciousness, guides attention which in turn guides the motor system in a way that leads to the successful execution of the movements required to steal the item. According to **S**, Bob's stealing the item was a voluntary action. Nevertheless, it seems strange to say that Bob is morally responsible—such that he can be justly *blamed*—for stealing the item, since he was coerced by fear of death to do it. So, voluntary action, as defined by **S**, is not sufficient to ground moral responsibility.

A somewhat *ad hoc* solution to this problem is just to say that an action for which one is morally responsible is one which is both *voluntary* (as defined by **S**) and *not coerced*. However, I think a more principled solution would be to add a *reflective endorsement* component to the account of voluntary action in **S** (Frankfurt, 1971). For example, if we add to this account of voluntary action the requirement that the brain *wants to have* some goal to act, then the account avoids the problem described above. In the scenario presented above, although Bob has the goal to steal the item, he does not *want to have this goal*. If wanting or identifying with or otherwise endorsing a goal is incorporated as a requirement in the account of voluntary action, then the account would predict that Bob's action was involuntary—a prediction which is in harmony with the intuition that Bob is not morally responsible for his action.

§9: Closing remarks

In this essay, I have reviewed a body of cognitive science literature about attention, consciousness, action, and the connections between these phenomena. I leveraged ideas and experimental results from this literature as support for two claims: (1) voluntary action requires voluntary attention and (2) voluntary attention requires consciousness. These claims jointly entail that voluntary action requires consciousness. I also reviewed a body of cognitive science literature about the subjective feeling of voluntariness. Lastly, I presented and critically examined a theoretical narrative that describes voluntary action as involving an interplay of goals, attention, consciousness, and movement.

References

Broughton *et al.* Homicidal Somnambulism: A Case Report. *Sleep* (1994); 17(3):253-64.

Baars, B. J. (1997). In the theatre of consciousness. Global Workspace Theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, *4*(4), 292–309. https://doi.org/10.1093/acprof:oso/9780195102659.001.1

Baars, B. J. (2002). The conscious access hypothesis:origins and recent evidence. *Trends in Cognitive Sciences*, *6613*(0), 47–52.

Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research*, *150*, 45–53. https://doi.org/10.1016/S0079-6123(05)50004-9

Bell, J. J., & Hardy, J. (2009). Effects of Attentional Focus on Skilled Performance in Golf. *Journal of Applied Sport Psychology*, *21*(2), 163–177. https://doi.org/10.1080/10413200902795323

Block, N. (1990). Consciousness and accessibility. *Behavioral and Brain Sciences*, *13*, 596–598.

Block, N. (1992). Begging the question against phenomenal consciousness. *Behavioral and Brain Sciences*, *15*, 205–206.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*, 227–247.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, *9*(2), 46–52. https://doi.org/10.1016/j.tics.2004.12.006

Borji, A., & Itti, L. (2013). State-of-the-art in Visual Attention Modeling. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, *35*(1), 185–207. https://doi.org/10.1109/TPAMI.2012.89

Camacho, E. F., & Bordons, C. (2007). *Model Predictive control*. *Journal of Chemical Information and Modeling*. https://doi.org/10.1007/978-0-85729-398-5

Camacho, E. F., & Bordons, C. (2007). Nonlinear Model Predictive Control: An Introductory Review. In *Assessment and Future Directions of Nonlinear Model Predictive Control* (Vol. 358, pp. 1–16). https://doi.org/10.1007/978-3-540-72699-9_1

Chambon, V., & Haggard, P. (2012). Sense of control depends on fluency of action selection, not motor performance. *Cognition*, *125*(3), 441–451. https://doi.org/10.1016/j.cognition.2012.07.011

Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2013). An online neural substrate for sense of agency. *Cerebral Cortex*, *23*(5), 1031–1037. https://doi.org/10.1093/cercor/bhs059

Chisholm, R. (1966). Freedom and Action. In K. Lehrer (Ed.), *Freedom and Determinism* (pp. 11–44). New York: Random House.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–15.

Dehaene, S., & Changeux, J. P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*, 193–222. Retrieved from http://www.annualreviews.org.proxy.lib.umich.edu/doi/pdf/10.1146/annurev.ne.18.030195.001205

Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, *18*(2), 254–261.

https://doi.org/10.1093/cercor/bhm050

Farrer, C., & Frith, C. D. (2002). Experiencing Oneself vs Another Person as Being the Cause of an Action: The Neural Correlates of the Experience of Agency. *NeuroImage*, *15*(3), 596–603. https://doi.org/10.1006/nimg.2001.1009

Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, *68*(1), 5–20.

Graziano, M. S. A., & Botvinick, M. M. (2002). How the brain represents the body: insights from neurophysiology and psychology. *Common Mechanisms in Perception and Action (Attention and Performance XIX)*. Retrieved from https://www.princeton.edu/~graziano/Papers/Attn_Perf19.pdf

Haggard, P., & Chambon, V. (2012). Sense of agency. *Current Biology*. https://doi.org/10.1016/j.cub.2012.02.040

Haggard, P., & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition*, *12*(4), 695–707. https://doi.org/10.1016/S1053-8100(03)00052-7

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat Neurosci*, *5*(4), 382–385. https://doi.org/10.1038/nn827\rnn827 [pii]

Haggard, P., & Cole, J. (2007). Intention, attention and the temporal experience of action. *Consciousness and Cognition*, *16*(2), 211–220. https://doi.org/10.1016/j.concog.2006.07.002

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research*, *126*(1), 128–133. https://doi.org/10.1007/s002210050722

Hartline, H. K., Wagner, H. G., & Ratliff, F. (1956). Inhibition in the eye of limulus. *The Journal of General Physiology*, *39*(5), 651–673.

Huxley, T. H. (1874). On the Hypothesis that Animals are Automata, and its History. *The Fortnightly Review*, *16*, 555–580.

Ivanoff, J., & Klein, R. M. (2003). Orienting of attention without awareness is affected by measurement-induced attentional control settings. *Journal of Vision*, *3*(1), 32–40. https://doi.org/10.1167/3.1.4

James, Wi. (1890). *The Principles of Psychology*. New York: Holt.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of Visual Attention in the Human Cortex. *Annual Review of Neuroscience*, *23*, 315–341.

Kelly, Y. T., Webb, T. W., Meier, J. D., Arcaro, M. J., & Graziano, M. S. a. (2014). Attributing awareness to oneself and to others. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(13), 5012–7. https://doi.org/10.1073/pnas.1401201111

Kentridge, R. W., Nijboer, T. C. W., & Heywood, C. A. (2008). Attended but unseen: Visual attention is not sufficient for visual awareness. *Neuropsychologia*, *46*(3), 864–869. https://doi.org/10.1016/j.neuropsychologia.2007.11.036

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. https://doi.org/10.1016/j.tics.2006.10.012

Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., Balasubramanian, V., & Sterling, P. (2006). How Much the Eye Tells the Brain. *Current Biology*, *16*(14), 1428–1434. https://doi.org/10.1016/j.cub.2006.05.056

Kuffler, S. W. (1953). Discharge Patterns and Functional of Mammalian Retina. *Journal of Neurophysiology*, *16*(1).

Lambert, A., Naikar, N., McLachlan, K., & Aitken, V. (1999). A new component of visual orienting: Implicit effects of peripheral information and subthreshold cues on covert attention. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(2), 321–340. https://doi.org/10.1037/0096-1523.25.2.321

Landau, A. N., Esterman, M., Robertson, L. C., Bentin, S., & Prinzmetal, W. (2007). Different effects of voluntary and involuntary attention on EEG activity in the gamma band. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(44), 11986–90. https://doi.org/10.1523/JNEUROSCI.3092-07.2007

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, *8*(4), 529–539. https://doi.org/10.1017/S0140525X00044903

Lohse, K. R., Jones, M., Healy, A. F., & Sherwood, D. E. (2014). The Role of Attention in Motor Control. *J Exp Pyschol Gen*, *143*(2), 930–48. https://doi.org/10.1017/CBO9781107415324.004

Lohse, K. R., & Sherwood, D. E. (2012). Thinking about muscles: The neuromuscular effects of attentional focus on accuracy and fatigue. *Acta Psychologica*, *140*(3), 236–245. https://doi.org/10.1016/j.actpsy.2012.05.009

Lohse, K. R., Sherwood, D. E., & Healy, A. F. (2010). How changing the focus of attention affects performance, kinematics, and electromyography in dart throwing. *Human Movement Science*, *29*(4), 542–555. https://doi.org/10.1016/j.humov.2010.05.001

Mack, A., & Rock, I. (1999). Inattentional Blindness An Overview By Arien Mack and Irvin Rock. *Psyche*, *5*. https://doi.org/10.1016/S1364-6613(98)01244-3

McCormick, P. a. (1997). Orienting attention without awareness. *Journal of Experimental*

Psychology. Human Perception and Performance, *23*(1), 168–180. https://doi.org/10.1037/0096-1523.23.1.168

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, 1–13. https://doi.org/10.1007/s00429-010-0262-0

Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, *18*(4), 1056–1064. https://doi.org/10.1016/j.concog.2009.05.004

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science (New York, N.Y.)*, *229*(4715), 782–784. https://doi.org/10.1126/science.4023713

Nagel, T. (1974). What Is It Like to Be A Bat? *The Philosophical Review*, *83*(4), 435–450. https://doi.org/10.2307/2183914

Norman, L. J., Heywood, C. a, & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychological Science*, *24*(6), 836–43. https://doi.org/10.1177/0956797612461449

Prinz, J. (2013). *The Conscious Brain*. *The Conscious Brain*. https://doi.org/10.1093/acprof:oso/9780195314595.001.0001

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(5), 1736–1753.

Scheidt, R. A., Conditt, M. A., Secco, E. L., & Mussa-Ivaldi, F. A. (2005). Interaction of visual and proprioceptive feedback during adaptation of human reaching movements. *Journal of*

Neurophysiology, *93*(6), 3200–13. https://doi.org/10.1152/jn.00947.2004

Schücker, L., Hagemann, N., Strauss, B., & Völker, K. (2009). The effect of attentional focus on running economy. *Journal of Sports Sciences*, *27*(12), 1241–1248. https://doi.org/10.1080/02640410903150467

Simard, T. G., & Basmajian, J. V. (1967). Methods in training the conscious control of motor units. *Archives of Physical Medicine and Rehabilitation*, *48*, 12–19.

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *NeuroImage*, *46*(1), 290–298. https://doi.org/10.1016/j.neuroimage.2009.01.047

Tsushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science (New York, N.Y.)*, *314*(5806), 1786–1788. https://doi.org/10.1126/science.1133197

Vance, J., Wulf, G., Töllner, T., McNevin, N., & Mercer, J. (2004). EMG activity as a function of the performer's focus of attention. *Journal of Motor Behavior*, *36*(4), 450–459. https://doi.org/10.3200/JMBR.36.4.450-459

Webb, T. W., & Graziano, M. S. A. (2015). The attention schema theory: a mechanistic account of subjective awareness. *Frontiers in Psychology*, *6*(April), 1–11. https://doi.org/10.3389/fpsyg.2015.00500

Webb, T. W., Igelström, K. M., Schurger, A., & Graziano, M. S. A. (2016). Cortical networks involved in visual awareness independent of visual attention. *Proceedings of the National Academy of Sciences*, *113*(48), 201611505. https://doi.org/10.1073/pnas.1611505113

Webb, T. W., Kean, H. H., & Graziano, M. S. A. (2016). Effects of Awareness on the Control of Attention. *Journal of Cognitive Neuroscience*, *28*(6), 842–851.

https://doi.org/10.1162/jocn_a_00931

Wegner, D. M. (2003). The mind's best trick: How we experience conscious will. *Trends in Cognitive Sciences*, *7*(2), 65–69. https://doi.org/10.1016/S1364-6613(03)00002-0

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882. https://doi.org/10.1126/science.7569931

Woollacott, M., & Shumway-Cook, A. (2002). Attention and the control of posture and gait: A review of an emerging area of research. *Gait and Posture*, *16*(1), 1–14. https://doi.org/10.1016/S0966-6362(01)00156-4

Wu, W. (2011). Attention as selection for Action. In *Attention: Philosophical and Psychological essays* (pp. 97–116).

Wu, W. (2016). Experts and Deviants: The Story of Agentive Control. *Philosophy and Phenomenological Research*, *93*(1), 101–126. https://doi.org/10.1111/phpr.12170

Wulf, G. (2007). Attentional Focus and Motor Learning: A review of 10 years of Research. *E-Journal: Movement and Training*, *1*(2–3), 1–11. https://doi.org/10.1080/1750984X.2012.723728

Wulf, G. (2013). Attentional focus and motor learning : A review of 15 years. *International Review of Sport and Exercise Psychology*, *6*(1), 77–104. https://doi.org/10.1080/1750984X.2012.723728

Wulf, G., Mcconnel, N., Gärtner, M., Schwarz, A., & Gartner, M. (2002). Enhancing the learning of sport skills through external-focus feedback. *Journal of Motor Behavior*, *34*(2), 171–82. https://doi.org/10.1080/00222890209601939

Wulf, G., & Su, J. (2007). An external focus of attention enhances golf shot accuracy in beginners and experts. *Research Quarterly for Exercise and Sport*, *78*(4), 384–389.

https://doi.org/10.5641/193250307X13082505158336

Yomogida, Y., Sugiura, M., Sassa, Y., Wakusawa, K., Sekiguchi, A., Fukushima, A., … Kawashima, R. (2010). The neural basis of agency: An fMRI study. *NeuroImage*, *50*(1), 198–207. https://doi.org/10.1016/j.neuroimage.2009.12.054

Zachry, T., Wulf, G., Mercer, J., & Bezodis, N. (2005). Increased movement accuracy and reduced EMG activity as the result of adopting an external focus of attention. *Brain Research Bulletin*, *67*(4), 304–309. https://doi.org/10.1016/j.brainresbull.2005.06.035