

CONSCIOUS AND UNCONSCIOUS MENTALITY

Examining their Nature, Similarities
and Differences

*Edited by Juraj Hvorecký, Tomáš Marvan and
Michal Polák*

First published 2024

ISBN: 978-1-032-52979-0 (hbk)

ISBN: 978-1-032-52974-5 (pbk)

ISBN: 978-1-003-40952-6 (ebk)

4

AGAINST UNCONSCIOUS VOLITION

Tim Bayne

(CC-BY-NC-ND) 4.0

DOI: 10.4324/9781003409526-5

This Chapter was funded by Canadian Institute for Advanced Research (CIFAR)



Routledge
Taylor & Francis Group
LONDON AND NEW YORK

4

AGAINST UNCONSCIOUS VOLITION

Tim Bayne

4.1 Introduction

In biology, a gonochoric species is one in which every individual is either male or female. Despite its rather unlovely sound, I will borrow this term to describe an influential—albeit rarely articulated—conception of mentality. According to what I will call ‘the gonochoric view of the mind’, mental categories are in general capable of manifesting as either conscious or unconscious. Just as bears, bats, and budgies can take either female or male form, so too beliefs, desires, intentions, perceptions and so on can take either conscious or unconscious form. That, at any rate, is the gonochoric picture of things.

Not everyone will treat this proposal as worthy of serious consideration. Those who hold that mental phenomena are, by definition, components of the stream of consciousness will dismiss it out of hand; so too will those who doubt that consciousness is a genuine phenomenon, albeit for very different reasons. But such views lie outside the mainstream, and most contemporary philosophers will take the gonochoric account seriously. Indeed, I suspect that the view is widely endorsed, although perhaps only in heavily qualified forms. At any rate, the unrestricted gonochoric account provides a useful stalking horse for anyone interested in the structure of mental categories and the relationship between mentality and consciousness.

In the recent literature, discussion of ‘gonochoric issues’ has centred on perception. Is (human) perception essentially conscious, or does it take both conscious and unconscious forms (see, e.g., Berger and Mylopoulos 2019; Block 2016; Phillips 2018; Phillips 2021; Phillips and Block 2016; Shepherd and Mylopoulos 2021)? Although this chapter is indebted to that debate, my focus here is on volition. I push back against leading currents in the philosophy

of agency to argue that a gonochoric conception of volition should be rejected, and that volition is essentially conscious.

The chapter unfolds as follows. Section 4.2 focuses on the gonochoric account, noting various difficulties which attend its formulation. Section 4.3 introduces the notion of volition (and volitional agency), and appeals to the anarchic hand syndrome to provide *prima facie* reason for treating unconscious volition with suspicion. Section 4.4 presents an argument ('the master argument') for the claim that volition is essentially conscious. Section 4.5 considers three objections to that claim: An objection from minimal actions and the exercise of expertise; an objection from the role of the readiness potential in initiating action; and an objection based on relational theories of consciousness. Section 4.6 concludes with some brief reflections on the gonochoric view of mentality and the nature of mental categories.

4.2 The gonochoric conception of mentality

The gonochoric conception of the mind is easy enough to grasp at an intuitive level, but providing a precise characterization of it is far from straightforward. Three clarifications are required.

The first issue concerns cognitive architecture. A gonochoric conception of category X holds that both conscious Xs and unconscious Xs are possible, but it does not imply that any agent that can X consciously will also be able to X unconsciously or vice-versa. Instead, the claim is only that the nature of Xs allows for both conscious Xs and unconscious Xs. Thus, even if both conscious and unconscious volitions are broadly possible, it might turn out that in certain types of cognitive architectures volition must be conscious, whereas in other types of cognitive architectures it cannot be conscious.

Second, we need to reckon with the personal/subpersonal contrast (Dennett 1969; Drayson 2014; Hornsby 2000). Personal-level phenomena are properly ascribed to the agent as such, whereas subpersonal phenomena characterize the operations of an agent's proper parts. Deciding to walk along a beach is a personal-level phenomenon (it's something that you do), but the computations that enable you to successfully navigate your way between the boulders on the beach are subpersonal. Any plausible version of the gonochoric account must be restricted to personal-level phenomena, for there is little reason to suppose that sub-personal states and processes can figure in consciousness.

Third, there are some personal-level mental categories for which a gonochoric treatment is clearly untenable, namely, those that are defined in terms of consciousness-involving properties. For example, the category of conscious states is clearly a mental category, but rejecting the gonochoric account on the grounds that conscious states are necessarily conscious would be a hollow victory indeed. What's needed here is some way of restricting the

gonochoric thesis to mental state types which are such that it's not obviously true that their instances either must be or cannot be conscious.

It's tempting to respond to this challenge by excluding the category of conscious states and its determinables from the scope of the states under discussion. Unfortunately, this suggestion is of limited help, for part of what's at issue in this debate is whether certain mental state terms refer to types of experiential phenomena. Consider the debate regarding the possibility of unconscious pain (Reuter and Sytsma 2020). Those who regard pain as necessarily conscious typically hold such a view because they take 'pain' to refer to a type of experience, whereas those who hold that unconscious pain is possible are likely to hold that view in part because they deny that 'pain' refers to a type of experience. In the case of biological species, it's relatively unproblematic (I take it) to identify the members of a single species without prejudging the question of whether or not that species is gonochoric. In the case of mentality, however, debate about whether mental particulars belong to the same kind (the same 'fundamental kind', if one likes) are not so easily disentangled from the question of whether that kind is gonochoric.

This challenge would need to be resolved by any attempt to address the gonochoric conception of mentality as such, but luckily it can be set to one side here. All that the present chapter requires is the assumption that the question can reasonably be raised as to whether volitions are gonochoric. And that assumption, it seems to me, is secure.

4.3 Volition and volitional agency

Talk of 'volition' may sound anachronistic, belonging more to the age of the typewriter and the steam engine than that of neuroimaging and the smartphone. However, although it may not be as central to the analysis of agency as it once was, the concept of volition provides a useful focus for thinking about the relationship between action and consciousness and remains a widely employed construct in both philosophical and empirical treatments of agency (e.g., Fried et al. 2011; Ginet 1990; Lowe 2008; Maoz and Sinnott-Armstrong 2021; Roskies 2010).

The term 'volition' is used in multiple ways. In one sense, it is a type of capacity. As Patrick Haggard writes, "volition is the capacity for endogenous action, particularly goal-directed endogenous action, shared by humans and some other animals" (Haggard 2019, 9). In this sense, then, we can describe a phenomenon as 'volitional' insofar as it involves an exercise of that capacity, just as we can describe a phenomenon as perceptual insofar as it involves the exercise of a perceptual capacity.

To a first approximation, volitions themselves can be thought of as representations that initiate and guide behaviour. (We will see that this characterization requires supplementation.) The contents of these

representations can straddle multiple levels of granularity, much in the way that perceptual contents can (see, e.g., Pacherie 2000; Siegel 2010). Some volitions will involve low-level goal-representations that (for example) describe the relationship between effectors and objects ('to move my arm thus-and-so in relation to this object'), whilst others will involve high-level goal-representations ('to pour a cup of tea').

Although some theorists might be tempted to identify volitions with volitional agency (and to hold that there is nothing more to acting than having a volition), I prefer to treat volitions as components of volitional agency. Consider William James's case of the individual who tries to raise his arm and thinks that he has done so, unaware that his arm hasn't moved due to the fact that it has been restrained. On certain conceptions of agency (e.g. Hornsby 1980), James has indeed acted, and the failure of his arm to move is due simply to the fact that his action does not have its intended effects. Another—and in my view more plausible—account holds that volitional agency requires both volition and the world's co-operation (as it were). In the same way that we need to distinguish perceptual experiences (which could be illusory or hallucinatory) from genuine perception, so too we should distinguish volitions (which could be thwarted) from genuine instances of volitional agency.

How do volitions differ from intentions? Intentions are standing propositional attitudes, and pattern with beliefs and desires insofar as each of these states can persist through periods of mental inactivity. Volitions, however, pattern with judgements and urges. They are occurrent, non-dispositional, phenomena—entities that exist only insofar as they actively impinge on thought and behaviour. Of course, volitions are not unrelated to intentions, just as judgements are not unrelated to beliefs and urges are not unrelated to desires. In order for intentions to effect change they must give rise to volitions, for it is volitions which initiate and guide behaviour.¹

What, then, of the relationship between volition and consciousness? Three views can be distinguished. First, there is the gonochoric view, according to which volitions can be either conscious or unconscious. This account has been defended by a number of authors, including Alfred Mele (2009) and David Rosenthal (2002). A second view treats volition as essentially unconscious. This view (or at least something very much like it) has been defended by Peter Carruthers (2007), although he uses the term 'willed agency' rather than 'volitional agency'.² Although there are important differences between Mele's and Rosenthal's accounts and that developed by Carruthers, both views are committed to the possibility of unconscious volition. That commitment is rejected by a third view, according to which volitions are essentially conscious. That account—which I take to reflect the traditional understanding of volition—is the one that I will defend here.³

In getting clear on some of the central issues here, it will be useful to consider the 'anarchic hand syndrome', also known as the 'alien hand syndrome'

(Della Sala et al. 1991; Feinberg et al. 1992; Gasquoine 1993; Goldberg and Bloom 1990; Marchetti and Della Sala 1998; Pacherie 2007). Associated with lesions of the corpus callosum and medial frontal damage, the anarchic hand syndrome is characterized by goal-directed agency on the part of one of the patient's hands (the 'anarchic' hand), which is at odds with the agent's wider aims and avowed intentions. The anarchic hand may spontaneously reach out and grasp door knobs (Goldberg et al. 1981); raise to the patient's lip a cup of tea that the patient knows is too hot to drink (Della Sala 1991); or take control of the steering wheel of the car that the patient is driving (Shao et al. 2019). Patients will attempt to restrain their anarchic ('bad') hand with their non-anarchic ('good') hand.

Do the movements of the anarchic hand involve unconscious volitions? In order not to prejudge the answer to that question, let's call the representations that initiate and guide anarchic behaviour 'goal representations' ('GRs'). To figure out whether GRs are unconscious volitions, we need to ask both whether they are volitions and whether they are unconscious. Let's tackle these questions in reverse order.

I assume here that a volition is (or would be) conscious in precisely the same sense in which other mental states are conscious—that is, it would be conscious in virtue of the fact that there is 'something distinctive that it's like' to instantiate it. Just as there's something it's like to (say) experience a sharp pain in one's leg, see a cat stroll along the top of a wall, or realize that a particular argument is invalid, so too there is something it's like to engage in volitional activity. Volitional phenomenology is a species of agentive phenomenology. Although ubiquitous, few would hold agentive phenomenology up as a paradigm instance of a conscious state, for it is rarely the explicit focus of consciousness. In picking up a glass, one's attention is typically focused on the glass, and one's awareness of one's own agency is at best a background component of the overall stream of consciousness. But although agentive experience is typically relegated to the margins of consciousness, its existence is introspectively detectable. (Or at least, so I would argue.)⁴

It is, of course, difficult to know what kinds of experiences (or lack thereof) anarchic hand patients have, but their behaviour suggests that they experience the movements of their anarchic hand in the way that you or I might experience the effects of an external agent. Patients will complain that the hand "will not do what I want it to do" (Goldberg et al. 1981), that it "has a mind of its own" (Feinberg et al. 1992), or that it "does what it wants to do" (Giovannetti et al. 2005). The patient is clearly not aware of 'their' anarchic actions in the way in which they are aware of their non-anarchic actions. Instead, their awareness of their anarchic actions appears to be akin to the awareness that one might have of another person's actions—awareness based on ordinary, perceptually mediated, inference.

If GRs are volitions, then they would appear to be unconscious volitions—but are they volitions at all? Thinking of GRs as volitions certainly has some plausibility. As we noted earlier, volitions initiate and guide behaviour, and GRs certainly appear to do that. The movements of the anarchic hand aren't genuinely 'anarchic', but are structured around goals of some kind. The hand (or is it the agent?) isn't merely moving, but is (e.g.) turning a doorknob, raising a cup of tea to the agent's lips, or moving a steering wheel. But for all that, it seems to me that instead of concluding that GRs are volitions, it would be more plausible to hold that volition involves more than the mere initiation and guidance of behaviour and deny that GRs are volitions.

The crucial point here is that anarchic agency is stimulus-driven rather than 'endogenous' (having its source 'within the agent'). Anarchic actions are triggered by perceptual engagement with a tool (a fork, a cup, a steering wheel) (Giovannetti et al. 2005; Kritikos et al. 2005), and are independent of the agent's intentions, commitments, beliefs, and so on. The agent knows that a certain door is not to be opened (but the hand reaches out for it nonetheless); the agent knows that the cup is too hot to drink (but the hand brings it to their lips); the agent knows that the car must be kept safe from oncoming traffic (but the hand turns the steering wheel in precisely that direction).⁵ In some sense anarchic actions can be compared to the kinds of actions that are produced by direct stimulation of a patient's motor cortex (e.g., Fried et al. 1991): It is the patient's body that moves, but 'the cause' of this movement is external to them. Intriguingly, clinical descriptions of the anarchic hand syndrome tend to ascribe the relevant actions to 'the hand' rather than to the patient, although of course treating hands as agents is far from unproblematic.⁶

Of course, ordinary behaviour is often elicited by perceptual affordances, too. One may have made a resolution to avoid sugary foods, but nonetheless 'find oneself' reaching for a doughnut at the office party. But there are significant contrasts between ordinary weak-willed behaviour and anarchic behaviour, for the sense in which one 'finds oneself' eating a doughnut (despite one's intentions/resolutions) is very different from the sense in which the anarchic hand patient 'finds themselves' moving a scalding cup of tea to their lips (despite their intentions/resolutions). For one thing, even akratic actions are experienced as having their source in oneself. Indeed, this is precisely what is so confronting about them. (One might feel less guilty about having the doughnut if one really did 'find oneself' reaching for it in the way in which the anarchic hand patient does.) Further, akratic actions have a kind of flexibility not seen in anarchic hand behaviour. The sight (and smell) of a ready-to-hand doughnut will activate the motor-programs appropriate for grasping it, but that activation can (and typically will) be inhibited when the stakes are high enough. I'll keep my hands in my pockets if you tell me that the doughnut is poisonous, or that those who resist doughnuts will be rewarded with one million dollars, or that a box of even tastier doughnuts is about to be

delivered. To the best of my knowledge, however, none of these interventions will have any effect on the ‘actions’ of the anarchic hand, and the patient will be able to prevent their anarchic hand from reaching for the doughnut only by restraining it in some way (e.g., by sitting on it).

Despite the goal-directed nature of anarchic behaviour, I conclude that it doesn’t involve volition (and hence doesn’t involve unconscious volition). But the upshot of these reflections isn’t purely negative, for they also suggest that there is more to the notion of volition than the mere initiation and guidance of behaviour. In order to give rise to volitional agency, goal representations must be properly ascribed to the agent—they must be suitably integrated with the agent’s intentions, commitments, beliefs, and so on. I turn now to explore this idea in more detail.

4.4 The master argument

The above reflections, I suggest, point to a more general line of argument against the possibility of unconscious volitions, according to which the possibility of unconscious volitions is ruled out on the grounds that volitions are personal-level phenomena. Unless a movement is guided by the agent as such then it’s not genuinely volitional, but at best ‘sub-volitional’, on a par with the kinds of sub-doxastic states that fall short of functioning as fully fledged judgements.

Of course, the notion of a personal-level process is clearly doing some heavy-lifting here, and there is no consensus account of what that notion involves. Here, I want to draw on Ian Phillips’s suggestion that the key to personal-level phenomenon is availability to ‘central coordinating agency’ (2018, 494). Unless a representation is able to directly contribute to central coordinating agency, it doesn’t qualify as personal-level. And directly contributing to central coordinating agency, I suggest, is possible only when the relevant representations are conscious. Taken together, these claims generate what I will call the master argument against the possibility of unconscious volition:

- (1) Volition is a personal-level phenomenon.
- (2) Personal-level phenomena must be able to contribute to central coordinating agency (CCA).
- (3) Only conscious representations can contribute towards central coordinating agency (CCA).

Therefore,

(C) Unconscious volitions are not possible.

I take (1) to be self-evident, and I won’t attempt to motivate it. No doubt some will reject it, but it seems to me as intuitively attractive—and as little in

need of defence—as the claim that perception is a personal-level phenomenon. (2) is perhaps not quite as uncontroversial as (1), but it too seems to me to be secure. Appealing to the notion of central coordinating agency is one of the two standard ways of drawing the personal/subpersonal contrast; indeed, it's really the only option that makes sense in this context given that the other way of drawing the contrast appeals to consciousness. (That conception of the subpersonal/personal contrast would, of course, secure the desired conclusion, but at the cost of making it dialectically toothless.)

What about (3)? Although certain conceptions of consciousness—for example, those associated with *a priori* versions of functionalism—might treat (3) as a conceptual truth, our concern here is with experiential (rather than functional) conceptions of consciousness. Thus, we can't simply help ourselves to (3) but must motivate it.

My preferred line of argument for (3) starts with perception rather than volition. As has often been noted (e.g., Shepherd 2015; Clark 2009), although unconscious sensory representations can facilitate and/or inhibit cognitive processing, if you want to bring information to bear on CCA—that is, if you want to bring it into rational contact with an agent's judgement, intentions and desires—then you need to make it consciously accessible to them. In Ned Block's (1995) useful phrase, phenomenality appears to 'grease the wheels' of cognitive accessibility. That thought, I suggest, applies not just to perception but also to volition: Unconscious representations cannot directly contribute to central coordinating agency, for the wheels of personal-level accessibility haven't been appropriately greased.

It is important to recognize that (3) does not entail that conscious representations *will* have the kind of personal-level integration that they ought to have. After all, failures of rational integration are a familiar feature of human psychology, as the phenomenon of action slips indicates (Amaya 2013; Mylopoulos 2021; Reason 1984). Consider the case of Pete, the distracted father. Although Pete normally drops his son (Oscar) off at kindergarten before he goes to work, today the kindergarten is closed, and childcare arrangements mean that Pete must look after Oscar at work. As Pete pulls into the parking lot his thoughts are preoccupied with work, and he leaves Oscar asleep in the back of the car. It's the middle of summer, and by the time that Pete realizes his mistake Oscar has succumbed to heat exhaustion.

There is nothing 'anarchic' about Pete's behaviour. He was aware of what he was doing as he parked his car and walked to his office. His problem wasn't that his behaviour was hijacked by motor representations of which he was unaware; rather, his problem is that his background propositional attitudes (his intention to look after his son; his knowledge that his son was in the back of the car) failed to generate the relevant kinds of occurrent representations when they should have. If they had done so, then Pete would surely not have left Oscar in the back of his car. Thus, action slips don't provide an objection to (3)—indeed, if anything they provide further reason to suppose

that consciousness is crucially involved in making representations available to central coordinating agency.

4.5 Objections and replies

4.5.1 *Minimal actions and the exercise of expertise*

As we have noted, agentic phenomenology rarely occupies the focus of consciousness. Indeed, it might be argued that there are everyday contexts in which one enjoys no agentic phenomenology at all—an obvious challenge to the position defended here.

Two kinds of cases need to be considered here. The first involve what Bach (1978) has called ‘minimal actions’. Commonplace examples would include pacing around a room whilst lost in thought, absent-mindedly doodling on a scratch-pad, or tapping one’s fingers on a desk whilst chatting to a friend. These actions don’t involve reflection or deliberation, but they do qualify as ‘doings’ rather than mere ‘happenings’.

One response to this challenge would be to treat minimal actions on the model of anarchic hand actions, and to suggest that they aren’t appropriately ascribed to the agent as such, but are instead triggered by features of the agent’s environment. But that, I think, wouldn’t be the right response. For one thing, minimal agency is not elicited by perceptual affordances in the way in which anarchic agency is. More importantly, minimal actions *are* appropriately sensitive to the agent’s knowledge and intentions in ways that anarchic hand behaviour isn’t. For example, one would surely stop pacing if one suddenly noticed a venomous snake in the corner. In this sense, then, minimal agency is appropriately ascribed to the agent in question whereas anarchic agency isn’t.

In my view, a better response to the challenge from minimal agency is to deny that such actions lack agentic phenomenology. It is certainly true that minimal actions tend not to grab our attention—indeed, it may even be a kind of conceptual truth that they must be relegated to the margins of awareness—but a mental phenomenon can characterize consciousness without falling within the focus of attention. If, during a period of time, one is absent-mindedly pacing around a room, one’s overall stream of consciousness will—I suggest—be characterized by a background awareness of this action.

A second challenge to the position developed here comes from the exercise of expertise, for it is sometimes claimed that experts have little to no awareness of the actions that they perform within their sphere of expertise (e.g., Beilock and Carr 2001; Dreyfus 2013; Brownstein 2014). Indeed, insofar as expertise is often associated with states of flow, it might be suggested that expertise not only doesn’t require volitional awareness but is positively undermined by it.

What might we say in response to this challenge? First, we need to acknowledge that experts retain ‘agent’s knowledge’ (Anscombe 1963/2000;

O'Brien 2003): experts know what they are doing without needing to explicitly attend to their performance. Here, of course, there is a marked contrast with anarchic hand patients, who lack 'agent's knowledge' of what 'they' are doing.

How might experts retain agent's knowledge if they lack conscious awareness of their volitions? The answer, I suggest, is that expertise does not in fact undermine agentic awareness—a point that has been made many times before (e.g., Sutton 2007; Montero 2010; Fridland 2014; Toner et al. 2014). Instead, the acquisition of expertise modulates the content and focus of agentic awareness. The novice must attend to the details of their environment and actions, whereas the expert is able to relegate much of what they do to the margins of awareness. What is thus relegated might leave little mark on long-term memory—consider David Armstrong's (1981) long-distance truck-driver, who suddenly 'comes to' and cannot recall their recent driving experience—but it may characterize the contents of awareness nonetheless. Instead of focusing on effector-based descriptors of what they are doing ('moving one's right leg thus-and-so'), the expert will typically prioritize larger, more complex, units of analysis (Christensen et al. 2016; Vallacher and Wegner 1987). For the expert, certain kinds of goal-directed representations become less salient; others take on heightened significance and require conscious monitoring.

What does this mean in terms of volition? Have we in effect granted that unconscious volitions are possible? Not quite. The suggestion is not that the development of expertise changes which volitions are conscious and which aren't, but that it changes which kinds of (sub-personal) motor representations qualify as (personal-level) volitions. In effect, expertise changes the 'shape' of agency, allowing the expert to chunk what they're doing into larger, differently shaped, units. The expert's volitions are no less conscious than the novice's—they just differ in scope.

4.5.2 *Volition and the readiness potential*

A second argument for the existence of unconscious volitions appeals to the readiness potential (RP, or *Bereitschaftspotential*) and, in particular, Libet's influential studies on the relationship between the RP and agentic awareness (Libet et al. 1983; Libet 1985).⁷ First described in 1965 by Kornhuber and Deecke, the RP refers to a slow build-up of scalp electrical potential preceding the onset of spontaneous movements. If, as Kornhuber and Deecke (1965) claimed, the RP functions as the "electro-physiological sign of planning, preparation, and initiation of volitional acts", then there is an interesting question as to how the RP is related to the agent's volitional awareness.

Libet set out to answer that question by instructing participants to perform a simple motor action (e.g., flexing their wrist) at the moment of their choosing within a specified period of time (say, 20 seconds), and to perform

this action ‘spontaneously’. They were also instructed to simultaneously monitor their agentic experiences, and to note the time at which they were first aware of the ‘decision’, ‘urge’ or ‘intention’ (terms that Libet used interchangeably) to move. Libet referred to the judgement of the time of their ‘decision’ (‘urge’, ‘intention’) to act as the ‘W judgement’. While subjects were both acting and monitoring their urges (intentions, decisions), Libet used an electroencephalogram (EEG) to measure the RP.⁸ Crucially, he found that although the RP preceded the movement by about 550 ms, participants appeared to be aware of their decision (urge, intention) only 200 ms prior to the action. In other words, there appeared to be a gap of about 350 ms between the RP and the point at which subjects claimed to be aware of their decision (urge, intention) to act.

Libet’s concern, of course, was with free will, and whether the relationship between the RP and the participants’ W judgements undermine human freedom. (Short answer: they don’t.) Our question, of course, is with the nature of volitions, and the question of whether volitions must be conscious. Do Libet’s data show that volitions can be unconscious?

Some commentators have certainly drawn that conclusion. For example, Mele (2009, 36) argues that the threat to free will can be defused by recognizing that proximal intentions (roughly, what I am here calling ‘volitions’) are not essentially conscious. Rosenthal (2002) makes a similar point, arguing that the RP is (or is the neural basis of) a volition that is initially unconscious but subsequently becomes conscious.

The apparent conflict with common sense, then, stems largely from adopting the view that a conscious state cannot also occur nonconsciously. Since we want to minimize the conflict between experimental results and common sense, we can see Libet’s results as providing evidence against that view.

(Rosenthal 2002, 218)

The interpretation of experimental results certainly should respect the commitments inherent in our folk psychological concepts, but it is far from clear that that constraint entails treating the RP as an unconscious volition. In fact, there are two other interpretations of Libet’s data, neither of which is committed to the existence of unconscious volitions.

The first of these accounts identifies the RP with an ‘urge’ or an ‘inclination’ to act rather than a volition as such. In defence of this proposal, Mele (2009) himself notes that the RP cannot be identified on particular trials, but is computed by back-averaging from the agent’s movement. Thus, he points out, we don’t know what percentage of RP trials are followed by action and what percentage of RP trials are not. And, if the RP is not typically followed by movement, the case for identifying it with volition is undermined. A second proposal has been defended at length by Aaron Schurger and his collaborators,

who have convincingly argued that the RP is an artifact of time-locking to a threshold-crossing event, and may not play any causally significant role in the production of movement (Schurger et al. 2012; Schurger et al. 2021). If this account is right, then the RP isn't associated with any kind of personal-level mental phenomenon at all. In short, there is ample reason to doubt that Libet's findings provide evidence of unconscious volition.

4.5.3 *Relational accounts of consciousness*

Rosenthal's case for unconscious volition appeals not only to Libet's studies, but also rests on an appeal to his relational conception of consciousness. According to Rosenthal, a mental state is conscious in virtue of the fact that the agent represents it in a particular manner (roughly, has an occurrent thought to the effect that they are in the corresponding state). Given this conception of consciousness, "[it] is never an essential property of anything that we are conscious of that thing" (2002, 218), as Rosenthal puts it. On the face of things, the claim that unconscious volitions are impossible would seem to be at odds with relational views of consciousness, and that in turn might be considered problematic given the influence of the relational approach to consciousness.

There are two things to say here. First, I would argue that the merits of the relational approach have often been oversold. Advocates of the view sometimes write as though its truth is conceptually guaranteed. Here is Rosenthal again:

No mental state counts as being conscious unless the individual who is in that state is conscious of the state. ... Since a state's being conscious consists of one's being conscious in a suitable way of being in that state, a state's being conscious cannot be an essential property of it. Any particular mental state can occur at one time consciously and at another not.

(Rosenthal 2002, 218; see also Lycan 2001)

However, the monitoring approach is by no means a trivial implication of the fact that 'a conscious state is a state that one is conscious of'. As Sosa (2003) has pointed out, the claim that 'one is conscious of one's conscious states' might be understood on the model of the claim that 'one smiles one's smiles' and 'one dances one's dances'. In other words, the crucial 'of' here might be the 'of' of possession and not the 'of' of intentionality.⁹ In my view, relational accounts are best viewed as scientific proposals, and not philosophical analyses whose truth might be ascertainable from the armchair.

The second point flows directly from the first, for those versions of the monitoring approach which are most directly informed by scientific considerations are invariably couched in subpersonal terms (see, e.g., Fleming 2020; Lau 2022). On these views, the relata of the consciousness-generating

relation are computational states rather than thoughts. But if consciousness emerges from relations between sub-personal representations, then there is no ‘straight line’ from relational accounts of consciousness to the conclusion that unconscious volition is possible. Any appearance to the contrary is the result of failing to properly distinguish personal from sub-personal levels of analysis.

4.6 Back to the gonochoric conception of mentality

Writing in relation to the unconscious perception debate, Berger and Mylopoulos (2019) make the following point:

[S]ince there would seem to be much common-sense and experimental evidence that *other* kinds of non-perceptual mental states—such as beliefs, desires, and emotions—can occur without being conscious, the strong sceptic [about unconscious perception] must either explain why perceptual states are unique in the mind in so far as they cannot occur unconsciously, or maintain that these other kinds of mental states cannot, despite appearances, occur unconsciously.

(2019, 25)

A parallel challenge can be mounted here: If other kinds of unconscious mental states—such as beliefs, desires, and emotions—are possible, why shouldn’t unconscious volitions also be possible? Call this the *uniformity challenge*, for it aims to put pressure on the claim that volitions are essentially conscious by suggesting that, if true, it would paint volitions as mysteriously unique.

Following Phillips (2021), the first point that must be made here is that mental phenomena do not belong to a single ontological category. Indeed, we have already contrasted standing mental states (such as intention, beliefs and desires) with occurrent (or episodic) phenomenon (such as perceptions, volitions, and bodily sensations). The idea that standing states are not necessarily conscious is not, it seems to me, overly controversial. Indeed, a case can be made for the claim that standing states *can’t* become conscious—that they don’t have the requisite structure (Crane 2013; Soteriou 2009). Whether or not that position is ultimately defensible, there are certainly grounds to deny that volitions ought to be treated in the same way as standing states when it comes to questions of consciousness.

But what about other occurrent mental phenomena, such as judgements, bodily sensations, and perceptual representations? Are these phenomena essentially conscious, or should we recognize the possibility of unconscious judgements, sensations, and perceptions? Addressing that question clearly goes well beyond the scope of this chapter. What can be said, however, is that it is far from obvious that treating volition as essentially conscious would flout any version of the uniformity principle that was tailored to occurrent,

personal-level, mental states. Indeed, one might even turn the uniformity challenge on its head, and argue that talk of unconscious volition ought to be regarded with as much skepticism as talk of unconscious judgement, unconscious pain, or indeed unconscious perception.

Acknowledgements

This publication was made possible in part through the support of a joint grant from the John Templeton Foundation and the Fetzer Institute (Consciousness and Free Will: A Joint Neuroscientific-Philosophical Investigation (John Templeton Foundation #61283; Fetzer Institute, Fetzer Memorial Trust #4189)). The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation or the Fetzer Institute. I thank Josh Shepherd and Liad Mudrik for discussion of the issues raised in this chapter, and am particularly grateful to Juraj Hvorecký and Michal Polák for their very helpful comments on an earlier draft.

Notes

- 1 Volitions, as I am conceiving of them here, correspond roughly to what John Searle (1983) has called intentions-in-action and what Alfred Mele (1992) has called ‘proximal intentions’. Searle distinguishes intentions-in-action from prior intentions in two ways: (1) some actions, such as spontaneously pacing around a room, occur without the formation of prior intentions, whereas all actions involve intentions-in-action; (2) prior intentions cause actions, but intentions-in-action are components of actions (Searle 1983, 84). Mele distinguishes proximal intentions from distal intentions, where a proximal intention is (roughly) a propensity to execute a plan for *immediate* action (where the plan is itself embedded in the intention), while a distal intention is a propensity to execute an intention-embedded plan for action in the *nonimmediate* future (Mele 1992, 144). For more on ‘intentions’ versus ‘volitions’ see Adams and Mele (1992).
- 2 Strictly speaking Carruthers doesn’t argue that volition (‘will’ in his terminology) is essentially unconscious, but only that the possibility of conscious will isn’t supported by the architecture of human cognition.
- 3 Other accounts of the relationship between agency and consciousness that are in keeping with the position developed here would include those developed by Bach (1978), Ginet (1990), Mossel (2005) and Searle (1983), although some of these authors avoid (indeed: disavow) talk of ‘volition’.
- 4 Agentive phenomenology is widely recognized as containing multiple strands, although there is dispute as to what those strands are and how they are related to each other. For reviews see: Bayne (2008); Horgan et al. (2003); Mylopolous and Shepherd (2020); and Pacherie (2008).
- 5 In this regard the behaviour of Dr. Strangelove in Stanley Kubrick’s *Dr Strangelove or: How I Learned to Stop Worrying and Love the Bomb* is somewhat misleading as a model of the anarchic hand syndrome, for Dr. Strangelove’s anarchic actions aren’t triggered by perceptually salient tools.

- 6 I suspect that the intuitive thought behind such descriptions is not that the hand itself is an agent but that anarchic behaviour involves ‘actions without an agent’, as Peacocke (2003) has suggested.
- 7 The literature on Libet’s studies is vast. For overviews see: Bayne 2011; Roskies 2011; Mele 2009; and Maoz et al. 2019.
- 8 Libet actually distinguished two types of readiness potentials. Spontaneous actions involve what he called a type II RP, whereas pre-planned actions exhibit what he called a type I RP, and can be found up to 1500 ms prior to the action (Libet et al. 1983; Tevena and Miller 2002).
- 9 Pigs and pains are both things of which one can be aware (or conscious), but it is far from clear that what it is to be a conscious pain (that is, a pain of which one is conscious) should be modelled on what it is to be a conscious pig (that is, a pig of which one is conscious).

References

- Adams, Frederick, and Alfred R. Mele. 1992. “The Intention/Volition Debate.” *Canadian Journal of Philosophy* 22 (3): 323–337. <https://doi.org/10.1080/00455091.1992.10717283>
- Amaya, Santiago. 2013. “Slips.” *Nous* 47 (3): 559–576. <https://doi.org/10.1111/j.1468-0068.2011.00838.x>
- Anscombe, Elizabeth G. 1963/2000. *Intention*. Cambridge, MA: Harvard University Press.
- Armstrong, David M. 1981. “The Nature of Mind.” In *The Nature of Mind and Other Essays*, edited by David Armstrong, 1–15. Ithaca, NY: Cornell University Press.
- Bach, Kent. 1978. “A Representational Theory of Action.” *Philosophical Studies* 34 (4): 361–379. <https://doi.org/10.1007/BF00364703>
- Bayne, Tim. 2008. “The Phenomenology of Agency.” *Philosophy Compass* 3 (1): 1–21. <https://doi.org/10.1111/j.1747-9991.2007.00122.x>
- Bayne, Tim. 2011. “Libet and the Case for Free Will Scepticism.” In *Free Will and Modern Science*, edited by Richard Swinburne, 25–46. Oxford: Oxford University Press.
- Beilock, Sian L., and Thomas H. Carr. 2001. “On the Fragility of Skilled Performance: What Governs Choking under Pressure?” *Journal of Experimental Psychology: General* 130 (4): 701–725. <https://doi.org/10.1037/0096-3445.130.4.701>
- Berger, Jacob, and Myrto Mylopoulos. 2019. “On Scepticism about Unconscious Perception.” *Journal of Consciousness Studies* 26 (11–12): 8–32.
- Block, Ned. 1995. “On a Confusion about a Function of Consciousness.” *Behavioral and Brain Sciences* 18 (2): 227–247. <https://doi.org/10.1017/S0140525X00038188>
- Block, Ned. 2016. “The Anna Karenina Principle and Scepticism about Unconscious Perception.” *Philosophy and Phenomenological Research* 93 (2): 452–459. <https://doi.org/10.1111/phpr.12258>
- Brownstein, Michael. 2014. “Rationalizing Flow: Agency in Skilled Unreflective Action.” *Philosophical Studies* 168 (2): 545–568. <https://doi.org/10.1007/s11098-013-0143-5>
- Carruthers, Peter. 2007. “The Illusion of Conscious Will.” *Synthese* 159 (2): 197–213. <https://doi.org/10.1007/s11229-007-9204-7>

- Christensen, Wayne, John Sutton, and Doris McIlwain. 2016. "Cognition in Skilled Action: Meshed Control and the Varieties of Skill Experience." *Mind and Language* 31 (1): 37–66. <https://doi.org/10.1111/mila.12094>
- Clark, Andy. 2009. "Perception, Action, and Experience: Unraveling the Golden Braid." *Neuropsychologia* 47 (6): 1460–1468. <https://doi.org/10.1016/j.neuropsychologia.2008.10.020>
- Crane, Tim. 2013. "Unconscious Belief and Conscious Thought." In *Phenomenal Intentionality*, edited by Uriah Kriegel, 156–173. Oxford: Oxford University Press.
- Della Sala, Sergio, Clelia Marchetti, and Hans Spinnler. 1991. "Right-sided Anarchic (alien) Hand: A Longitudinal Study." *Neuropsychologia* 29 (11): 1113–1127. [https://doi.org/10.1016/0028-3932\(91\)90081-i](https://doi.org/10.1016/0028-3932(91)90081-i)
- Dennett, Daniel. 1969. *Consciousness and Content*. Routledge.
- Drayson, Zoe. 2014. "The Personal/Subpersonal Distinction." *Philosophy Compass* 9 (5): 338–346. <https://doi.org/10.1111/phc3.12124>
- Dreyfus, Hubert L. 2013. "The Myth of the Pervasiveness of the Mental." In *Mind, Reason, and Being-in-the-World*, edited by Joseph K. Schear, 15–38. London: Routledge.
- Feinberg, Todd E., Rachel J. Schindler, Natalie G. Flanagan, and Laurence D. Haber. 1992. "Two Alien Hand Syndromes." *Neurology* 42 (1): 19–24. <https://doi.org/10.1212/wnl.42.1.19>
- Fleming, Stephen M. 2020. "Awareness as Inference in a Higher-Order State Space." *Neuroscience of Consciousness* 2020 (1): niz020. <https://doi.org/10.1093/nc/niaa011>
- Fridland, Ellen. 2014. "They've Lost Control: Reflections on Skill." *Synthese* 191 (12): 2729–2750. <https://doi.org/10.1007/s11229-014-0411-8>
- Fried, Itzhak, Amiram Katz, Gregory McCarthy, Kimberlee J. Sass, Peter Williamson, Susan S. Spencer, and Dennis D. Spencer. 1991. "Functional Organization of Human Supplementary Motor Cortex Studied by Electrical Stimulation." *Journal of Neuroscience* 11 (11): 3656–3666. <https://doi.org/10.1523/JNEUROSCI.11-11-03656.1991>
- Fried, Itzhak, Roy Mukamel, and Gabriel Kreiman. 2011. "Internally Generated Preactivation of Single Neurons in Human Medial Frontal Cortex Predicts Volition." *Neuron* 69 (3): 548–562. <https://doi.org/10.1016/j.neuron.2010.11.045>
- Gasquoine, Philip G. 1993. "Alien Hand Sign." *Journal of Clinical and Experimental Neuropsychology* 15 (5): 653–667. <https://doi.org/10.1080/01688639308402587>
- Ginet, Carl. 1990. *On Action*. Cambridge: Cambridge University Press.
- Giovannetti, Tania, Laurel J. Buxbaum, Iftah Biran, and Anjan Chatterjee. 2005. "Reduced Endogenous Control in Alien Hand Syndrome: Evidence from Naturalistic Action." *Neuropsychologia* 43 (1): 75–88. <https://doi.org/10.1016/j.neuropsychologia.2004.06.017>
- Goldberg, Gary, and Karen K. Bloom. 1990. "The Alien Hand Sign. Localization, Lateralization, and Recovery." *American Journal of Physical Medicine and Rehabilitation* 69 (5): 228–238. <https://doi.org/10.1097/00002060-199010000-00002>
- Goldberg, Gary, Nathaniel H. Mayer, and Joseph U. Togli. 1981. "Medial Frontal Cortex Infarction and the Alien Hand Sign." *Archives of Neurology* 38 (11): 683–688. <https://doi.org/10.1001/archneur.1981.00510110043004>

- Haggard, Patrick. 2019. "The Neurocognitive Bases of Human Volition." *Annual Review of Psychology* 70: 9–28. <https://doi.org/10.1146/annurev-psych-010418-103348>
- Horgan, Terence, John L. Tienson, and George Graham. 2003. "The Phenomenology of First-Person Agency." In *Physicalism and Mental Causation*, edited by Sven Walter and Heinz-Dieter Heckmann, 323–340. Exeter, UK: Imprint Academic.
- Hornsby, Jennifer. 1980. *Actions*. London: Routledge.
- Hornsby, J. 2000. "Personal and Sub-Personal: A Defense of Dennett's Early Distinction." *Philosophical Explorations* 3 (1): 6–24. <https://doi.org/10.1080/13869790008520978>
- Kornhuber, Hans H., and Lüder Deecke. 1965. "Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale." *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere* 284: 1–17. <https://doi.org/10.1007/BF00412364>
- Kritikos, Ada, Nora Breen, and Jason B. Mattingley. 2005. "Anarchic Hand Syndrome: Bimanual Coordination and Sensitivity to Irrelevant Information in Unimanual Reaches." *Cognitive Brain Research* 24 (3): 634–647. <https://doi.org/10.1016/j.cogbrainres.2005.03.015>
- Lau, Hakwan. 2022. *In Consciousness We Trust*. Oxford: Oxford University Press.
- Libet, Benjamin. 1985. "Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action." *Behavioral and Brain Sciences* 8 (4): 529–566. <https://doi.org/10.1017/S0140525X00044903>
- Libet, Benjamin, Curtis A. Gleason, Elwood W. Wright, and Dennis K. Pearl. 1983. "Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential). The Unconscious Initiation of a Freely Voluntary Act." *Brain* 106 (Pt 3): 623–642. <https://doi.org/10.1093/brain/106.3.623>
- Lowe, Edward J. 2008. *Personal Agency*. Oxford: Oxford University Press.
- Lycan, William. 2001. "A Simple Argument for a Higher-Order Representation Theory of Consciousness." *Analysis* 61 (269): 3–4. <https://doi.org/10.1111/1467-8284.00261>
- Maoz, Uri, and Walter Sinnott-Armstrong (eds.). 2021. *Free Will: Philosophers and Neuroscientists in Conversation*. Oxford: Oxford University Press.
- Maoz, Uri, Gideon Yaffe, Christof Koch, and Liad Mudrik. 2019. "Neural Precursors of Decisions that Matter—an ERP Study of Deliberate and Arbitrary Choice." *eLife* 8: e39787. <https://doi.org/10.7554/eLife.39787>
- Marchetti, Clelia, and Sergio Della Sala. 1998. "Disentangling the Alien and Anarchic Hand." *Cognitive Neuropsychiatry* 3 (3): 191–207. <https://doi.org/10.1080/135468098396143>
- Mele, Alfred R. 1992. *Springs of Action*. Oxford: Oxford University Press.
- Mele, Alfred R. 2009. *Effective Intentions*. Oxford: Oxford University Press.
- Montero, Barbara. 2010. "Does Bodily Awareness Interfere with Highly Skilled Movement?" *Inquiry* 53 (2): 105–122. <https://doi.org/10.1080/00201741003612138>
- Mossel, Benjamin. 2005. "Actions, Control and Sensations of Acting." *Philosophical Studies* 124 (2): 129–180. <https://doi.org/10.1007/s11098-004-7816-z>
- Mylopoulos, Myrto. 2021. "Oops! I Did it Again: The Psychology of Everyday Action Slips." *Topics in Cognitive Science* 14 (2): 282–294. <https://doi.org/10.1111/tops.12552>

- Mylopoulos, Myrto, and Joshua Shepherd. 2020. "The Experience of Agency." In *Oxford Handbook of the Philosophy of Consciousness*, edited by Uriah Kriegel, 164–187. Oxford: Oxford University Press.
- O'Brien, Lucy. 2003. "On Knowing One's Own Actions." In *Agency and Self-Awareness*, edited by Naomi Eilan and Johannes Roessler, 358–382. Oxford: Clarendon Press.
- Pacherie, Elisabeth. 2000. "Levels of Perceptual Content." *Philosophical Studies* 100 (3): 237–254. <https://doi.org/10.1023/A:1018624608680>
- Pacherie, Elisabeth. 2007. "The Anarchic Hand Syndrome and Utilization Behavior: A Window onto Agentive Self-awareness." *Functional Neurology* 22 (4): 211–217.
- Pacherie, Elisabeth. 2008. "The Phenomenology of Action: A Conceptual Framework." *Cognition* 107 (1): 179–217. <https://doi.org/10.1016/j.cognition.2007.09.003>
- Peacocke, Christopher. 2003. "Awareness, Ownership, and Knowledge." In *Agency and Self-Awareness*, edited by Naomi Eilan and Johannes Roessler, 94–110. Oxford: Clarendon Press.
- Phillips, Ian. 2018. "Unconscious Perception Reconsidered." *Analytic Philosophy* 59 (4): 471–514. <https://doi.org/10.1111/phib.12135>
- Phillips, Ian. 2021. "Scepticism about Unconscious Perception Is the Default Hypothesis." *Journal of Consciousness Studies* 28 (3-4): 186–205.
- Phillips, Ian, and Ned Block. 2016. "Debate on Unconscious Perception." In *Current Controversies in Philosophy of Perception*, edited by Bence Nanay, 165–192. New York: Routledge.
- Reason, James. 1984. "Little Slips and Big Disasters." *Interdisciplinary Science Reviews* 9 (2): 179–189. <https://doi.org/10.1179/isr.1984.9.2.179>
- Reuter, Kevin, and Justin Sytsma. 2020. "Unfelt Pain." *Synthese* 197 (4): 1777–1801. <https://doi.org/10.1007/s11229-018-1770-3>
- Rosenthal, David. 2002. "The Timing of Conscious States." *Consciousness and Cognition* 11 (2): 215–220. <https://doi.org/10.1006/ccog.2002.0558>
- Roskies, Adina L. 2010. "How Does Neuroscience Affect Our Conception of Volition?" *Annual Review of Neuroscience* 33: 109–130. <https://doi.org/10.1146/annurev-neuro-060909-153151>
- Roskies, Adina L. 2011. "Why Libet's Studies Don't Pose a Threat to Free Will." In *Conscious Will and Responsibility*, edited by Walter Sinnott-Armstrong and Lynn Nadel, 11–22. New York: Oxford University Press.
- Schurger, Aaron, Pengbo B. Hu, Joanna Pak, and Adina L. Roskies. 2021. "What is the Readiness Potential?" *Trends in Cognitive Sciences* 25 (7): 558–570. <https://doi.org/10.1016/j.tics.2021.04.001>
- Schurger, Aaron, Jacobo D. Sitt, and Stanislas Dehaene. 2012. "An Accumulator Model for Spontaneous Neural Activity Prior to Self-initiated Movements." *Proceedings of the National Academy of Sciences of the United States of America* 109 (42): E2904–E2913. <https://doi.org/10.1073/pnas.1210467109>
- Searle, John. 1983. *Intentionality*. Cambridge: Cambridge University Press.
- Shao, Jie, Rongrong Bai, Guangyu Duan, Yujia Guan, Li Cui, and Hui Deng. 2019. "Intermittent Alien Hand Syndrome Caused by Marchiafava-Bignami Disease: A Case Report." *Medicine* 98 (34): 1–5. <https://doi.org/10.1097/MD.00000000000016891>
- Shepherd, Joshua. 2015. "Conscious Control Over Action." *Mind and Language* 30 (3): 320–344. <https://doi.org/10.1111/mila.12082>

- Shepherd, Joshua, and Myrto Mylopoulos. 2021. "Unconscious Perception and Central Coordinating Agency." *Philosophical Studies* 178: 3869–3893. <https://doi.org/10.1007/s11098-021-01629-w>
- Siegel, Susanna. 2010. *The Contents of Visual Experience*. Oxford: Oxford University Press.
- Sosa, Ernest. 2003. "Privileged Access." In *Consciousness: New Philosophical Perspectives*, edited by Quentin Smith and Aleksander Jokić, 238–251. Oxford: Oxford University Press.
- Soteriou, Matthew J. 2009. "Mental Agency, Conscious Thinking, and Phenomenal Character." In *Mental Actions*, edited by Lucy O'Brien and Matthew J. Soteriou, 231–252. Oxford: Oxford University Press.
- Sutton, John. 2007. "Batting, Habit and Memory: The Embodied Mind and the Nature of Skill." *Sport in Society* 10 (5): 763–786. <https://doi.org/10.1080/17430430701442462>
- Toner, John, Barbara G. Montero, and Aidan Moran. 2014. "Considering the Role of Cognitive Control in Expert Performance." *Phenomenology and the Cognitive Sciences* 14 (4): 1127–1144. <https://doi.org/10.1007/s11097-014-9407-6>
- Trevena, Judy A., and Jeff Miller. 2002. "Cortical Movement Preparation before and after a Conscious Decision to Move." *Consciousness and Cognition* 11 (2): 162–190. <https://doi.org/10.1006/ccog.2002.0548>
- Vallacher, Robin R., and Wegner, Daniel M. 1987. "What Do People Think They're Doing? Action Identification and Human Behavior." *Psychological Review* 94 (1): 3–15. <https://doi.org/10.1037/0033-295X.94.1.3>