A predictive processing theory of motivation

Alex James Miller Tate¹

Received: 11 April 2019 / Accepted: 31 July 2019 / Published online: 10 October 2019
© The Author(s) 2019

Abstract

In this paper I propose minimal criteria for a successful theory of the mechanisms of motivation (i.e. how motivational mental states perform their characteristic function), and argue that extant philosophical accounts fail to meet them. Further, I argue that a predictive processing (PP) framework gives us the theoretical power to meet these criteria, and thus ought to be preferred over existing theories. The argument proceeds as follows—motivational mental states are generally understood as mental states with the power to *initiate*, *guide*, and *control* action, though few existing theories of motivation explicitly detail how they are meant to explain these functions. I survey two contemporary theories of motivational mental states, due to Wayne Wu and Bence Nanay, and argue that they fail to satisfactorily explain one or more of these functions. Nevertheless, I argue that together, they are capable of giving a strong account of the control function, which competing theories ought to preserve (all else being equal). I then go on to argue that what I call the ‘predictive theory’ of motivational mental states, which makes use of the notion of *active inference*, is able to explain all three of the key functions and preserves the central insights of Wu and Nanay on control. It thus represents a significant step forward in the contemporary debate.

Keywords Motivation · Predictive processing · Action · Control · Pragmatic representations · Attention

1 Introduction

The main goal of this paper is to develop and defend a theory of the cognitive mechanisms underlying the operation of motivational mental states. I aim for it to provide a plausible explanation of the key functions of motivation that have previously been identified in the literature—namely motivation’s role in the **initiation**, **guidance**, and

✉ Alex James Miller Tate
alexander.miller_tate@kcl.ac.uk

¹ The Dickson Poon School of Law, Somerset House, King’s College London, Strand, London, UK

control of action¹ (Shepherd 2017; Pacherie 2012: p. 96; Elliott 2006). I shall begin by clarifying the notion of motivation I am working with, before moving on to clarifying the question this paper will concern itself with.

When philosophers and psychologists speak of motivation, they typically speak of *motivational mental states*. These are typically glossed as any mental state, or state-complex, with the power to generate action (I will clarify this shortly). The paradigm of this class of mental states is an *intention*, but the class more generally is widely thought to include states such as emotions (Scarantino and Nielsen 2015; Tappolet 2010; Frijda 1986), pains (Corns 2014; Bain 2011), desires (Smith 1994) or action-desires (Mele 1992), and perhaps even certain kinds of belief² (McDowell 1979; Dancy 1993).

Further, many action theorists also distinguish between *proximal* and *distal* intentions³ (Pacherie 2012; Mele 1992; Bratman 1987; Searle 1983). Proximal intentions are those that have the power to initiate, guide, and control action (Shepherd 2017; Pacherie 2012: p. 96). Traditionally, their playing these three functions at least partially explains how they (at least according to a *causal* theory of action, which I shall assume throughout⁴) are able to cause bodily behaviours in such a way that they count as actions rather than mere reflexes. If I pick up a glass of water, then that bodily behaviour counts as an action only if it was caused (non-deviantly⁵) by an intention of mine to pick up the glass. The way to think about such intentions is as standing states (Wu 2011) that are disposed to act as the immediate mental antecedents (Nanay 2013) of particular bodily actions, when suitable circumstances arise or, at least, are perceived or judged to have arisen.

A distal intention is one that is associated with the functions intentions have been suggested to play in practical reasoning (see Pacherie 2012: pp. 95–96; Bratman 1987). That is, they are those kinds of intentions that act as prompters and terminators of different forms of practical reasoning ('means-end' and 'what-to-do' reasoning respectively), as well as coordinating agents' activities over time and with other agents.

The labels of 'proximal' and 'distal' are thus roughly intended to capture a distinction between the motor and deliberative functions of intention. Some also think that this distinction does some significant metaphysical work in action theory proper,

¹ Not all elements of the literature identify these functions in quite the same way—Elliott (2006) for instance speaks of motivation as that which 'energises' and 'directs' behaviour. This does not reflect a significantly different opinion regarding motivation's core functions—it is simply a less precise characterisation of the same basic functions mentioned above.

² I use 'belief' here as a convenient shorthand; what McDowell and Dancy think is *actually* motivating is some kind of 'moral conception/perception of the situation', which may or may not count as a belief per se.

³ Terminology here is far from uniform—Searle, for instance, distinguishes between *intentions-in-action* and *prior intentions*, while having basically the same distinction in mind (1983). Here, I use Mele's (1992) terminology.

⁴ Typically, these theories of action contrast with *agent causal theories*, which posit that only entire agents, rather than any individual state (or state-complex) of those agents, can cause bodily behaviours in such a way that they count as actions (see O'Connor 2011 for an overview). I will not consider such theories here, preferring to address my remarks towards those broadly located in the traditional causal camp.

⁵ What non-deviance exactly consists in is an open question that I will not try to answer here—see Wu (2016) and Shepherd (forthcoming) for some recent suggestions.

for instance by opening up a path to a solution of the problem of minimal actions⁶ (Pacherie 2012: p. 96; Searle 1983).

My focus in this paper will be on motivational mental states, understood as mental states with the power to play the roles ascribed above to proximal intentions; action initiation, guidance, and control (I will unpack these notions in the next section). Being able to cause action, rather than being able to play a particular kind of deliberative role is, in my view, the basic power that distinguishes motivational mental states from motivationally inert ones.

One reason to think this is that it is widely agreed that emotions are motivational mental states. But emotions are not always (perhaps even often) thought of as being the initiators or terminal points of practical reasoning, but rather as being directly elicited by environmental triggers or stimuli, such as threats or losses. What they *can* intuitively do, however, is cause actions; the fear I experience when I see a spider can intuitively cause me to do things that are far more than simple reflexes—I might call for my housemate if I think she is in, or cautiously retreat from the room and seek out a mug and a postcard. Absent other criteria, it thus makes sense to characterise both emotions and intentions as motivational mental states in virtue of their action-causing power and function, since this is what they have in common, at least on the surface.

Another reason; it is clearly this direct *action-causing* power that is at stake in philosophical debates over what sort of mental states are able to count as motivational (for instance between Humeans and Anti-Humeans). Thus, while Michael Smith (1994) thinks that beliefs are never *by themselves* motivational (always requiring the presence of some attendant desire, or ‘pro-attitude’ to actually generate action), Jonathan Dancy (1993) disagrees. Properly understood, according to him, our moral beliefs alone are sufficient to move us to act. What is at stake in their exchange is precisely whether (some) moral beliefs are able to cause action without an associated desire, in the way Dancy claims and Smith denies. Thus, we see that moral belief’s status as a motivational mental state-type hangs on its purported capacity to directly cause action. And the functions outlined above are simply a breakdown of the functional roles mental states are thought to need to play in order to count as *causing action*—they must (at least) initiate, guide, and control bodily behaviour.

More exactly, my focus in this paper is the mechanism by which motivational mental states play their action-initiating, guiding, and controlling functions. Why do I focus on these three?

Primarily, the choice is a pragmatic one, informed by the facts that, (1) these are the functions most often identified in the wider literature as under the purview of motivational mental states, and (2) they are thought to be the key components of motivational mental states’ overarching action-causing function (Shepherd 2017; Pacherie 2012). Since identifying *all* relevant functional roles of motivational mental states in action

⁶ Minimal actions are those which are performed routinely, unthinkingly, and pre-attentively. Examples include shifting gear while driving or pacing anxiously across a room (Searle 1983). The problem is that while such bodily behaviours seem much too complex and controlled to count as mere reflexes, they are apparently not preceded by a clear, specific intention of which the agent is aware. So how can a causal theory class them as actions (Bach 1978)? The standard strategy is for a causal theorist to claim that while such actions are not preceded by *distal* intentions (which are presumed to be the kind of which an agent are largely aware) they *are* nevertheless preceded by some kind of *proximal* intention. The details and difficulties of various specific responses of this kind need not concern us here.

appears to be a significant ongoing empirical and conceptual task,⁷ we must begin somewhere.

Assuming that theorists are correct about these three distinguishing functional roles of motivational mental states, it is clearly a necessary feature of any theory of motivation that it be able to explain how they are performed. It may well be that there are other critical desiderata; perhaps the ability to explain other critical functional roles yet to be widely identified, for instance. But it is important to recognise that if a purported theory of motivational mechanisms does not clearly specify how motivational mental states can initiate, guide, and control behaviour, then it is failing to explain the functional roles that distinguish motivational states from motivationally inert states. Thus, it cannot qualify as an adequate theory of motivation. Naturally, it might count as a *partial* theory of motivation, if it successfully accounts for some but not all of these central functional roles or takes incomplete steps toward accounting for them. Nevertheless, a minimal adequacy condition on a *complete* theory of motivation is that it account for all three of these central functional roles.

A concern one might have at this point is that it appears as if this whole enterprise is founded on the following naïve suggestion; a theory of motivation can only succeed if it shows how a *single* mental state (typically, a proximal intention or similar) can play the roles of action initiation, guidance, and control. And it's far from obvious that such a supposition is reasonable; though we might begin our theorising with a folk-psychological notion of intention, there may turn out to be no one-to-one mapping between this category and the computational/neural components of motivational architecture.⁸

Certainly, this concern is well-placed. We should not build this kind of assumption into our theorising about motivation from the beginning. So, I will not; a theory of motivation may posit as many kinds of mental state or state-complexes that it needs in order to explain the central functional roles of motivation, as long as this positing of extra states plays some demonstrable explanatory role. I take this last condition to be a fairly weak requirement of quantitative parsimony (see Nolan 1997).

That said, it is worth noting that whether we see motivation as being dealt with by one or more states probably depends more on the level of explanation at which we work than any objective ontological fact. Functionally speaking, motivational mental states, as we are working with the term here, *just are* the states or combinations of states that play the functional roles of initiation, guidance, and control (and perhaps some others). There is no particular reason why we should worry about whether *in fact* these roles are fundamentally played by a single state or many. Consequently, I will speak throughout as if all the work of motivational mental states is done by single states, not many. If the reader objects to this kind of simplification, then I invite them to read my claims as implicitly referring to many states, or complexes of several states, or similar. Nothing I shall have to say hangs on reading these claims strictly in the singular.

⁷ Recently, for instance, Wulf et al. (2010) provide evidence that motivational mental states play a crucial role in processes of complex motor skill learning.

⁸ I thank an anonymous reviewer for *Synthese* for raising this worry.

So, we have some sense of what a theory of motivation ought to explain. But the main work of the paper is, naturally, still ahead of us. In Sect. 2, I unpack the notions of initiation, guidance, and control in more detail, providing a clear explanatory target to guide the rest of the paper. After that, in Sect. 3, I discuss two extant theories of motivation (due to Wayne Wu and Bence Nanay), and identify how each falls short of meeting the explanatory targets identified in Sect. 2. I do, however, note that each theory provides helpful insights that ought to be preserved in any alternative theory. Next, in Sect. 4, I introduce the predictive processing framework that will form the core of my positive proposal, including the critical notion of **active inference**. Finally, in Sect. 5, I use this framework to present an alternative theory of motivation; one which can account for how motivational mental states fulfil their three key functions.

My central claim is that motivational mental states play their role of initiating, guiding and controlling action by causing the prediction of, and selective redeployment of attention towards, action-relevant proprioceptive and exteroceptive sensory signals. The devil, naturally, is in the details, which I shall explain closely in Sect. 5.

2 Three functional roles for motivation

This section unpacks the three central functions of motivational mental states identified above, providing a clear explanatory target for the rest of the paper.

2.1 Initiation

Let's start with action initiation. This is the most basic of the functions ascribed to motivational mental states by causal theories of action. On any such theory, motivational mental states are “responsible for triggering or initiating the intended action” (Pacherie 2012: p. 96). No causal theory of action would be worthy of the name if it did not claim that motivational mental states of some kind or another possess this function, simply because whatever else intentions do for processes of action, if they do not trigger them, they cannot be sensibly said to *cause* them.

Notice that it would be insufficient for a theory to suggest that motivational mental states somehow *simplify* the process of initiating action, or that they merely *enable* action, or that they provide some *background conditions* necessary for the action to be initiated. None of these things by themselves amount to anything like *causing* the relevant action. Simplifying, enabling, or setting up some necessary background conditions for action to be initiated are probably necessary functions of motivational mental states in causing actions, but they are not by themselves sufficient. For a theory of motivation to account for the initiating function, it needs to posit a mechanism by which motivational mental states play a *direct*, rather than merely *supportive* role in causing actions.

To help distinguish the notion of a direct as opposed to a supporting role in causing action, consider the following. If I successfully climb a ladder, it is plausible that *many* different kinds of mental states will have supported my doing so—my belief that there is a ladder to be found somewhere in the environment, visual states that inform me of

the actual presence of a ladder, its location, shape, and so on, proprioceptive states that inform me of the relative position of my hands and feet, etc. It is possible that without any of these mental states/processes I would fail to successfully climb the ladder. But when I do successfully climb a ladder, it is unnatural to say that they alone initiated the action. Rather, they are enablers of my climbing. On the other hand, it is considerably more natural to speak of my *intention* to climb the ladder as the thing that initiated my climbing behaviour (even though it is necessarily supported by a range of other states). Even though I will likely not climb a ladder (even if I intend to) if I cannot find one, it seems wrong, or at least unnatural, to answer a question of ‘Why did you climb that ladder?’ with ‘Because I saw it’. After all, I regularly see ladders and do not climb them. What makes this kind of distinction meaningful, to my mind, is at least in part the fact that intentions (or your preferred alternative motivational mental state/complex) seem to be the *difference-makers* in the cases where I *do* climb ladders. That is, what distinguishes between cases where I see ladders and do climb them and cases where I see ladders and do *not* climb them is plausibly an intention to climb the ladder in question.

Note that this thought holds even if one thinks that perceptual states are necessary components of motivational state-complexes. Since perceptions of objects can occur in the absence of any motivation to act on them in any way, it is implausible to suppose that these kinds of states explain the initiation of action, even when they are a necessary part of state-complexes, some component of which *does* explain the initiation of action.

It is important that theories of motivation have a principled explanation of this intuitive distinction—they must explain why it is right to think of motivational states (rather than any other kind of state) as the key difference makers in cases where we do act, as contrasted with cases where we do not; glibly, how motivational mental states actually make actions happen, rather than how they make it *merely possible* for actions to happen.

An immediate suggestion regarding what makes this difference between motivation inertness and motivational activity is the *causal proximity* of the state in question to the relevant bodily behaviour. That is, one might suggest that a motivational mental state *initiates* a behaviour just in case it is the *maximally proximal* psychological cause of the action in question. The reason why visual states are motivationally inert (and merely support action) is that they are too causally downstream of any actual behaviour. The bet, then, would be that proximal intentions are the maximally proximal element of a causal chain that terminates in bodily behaviour.

This sort of suggestion would be a mistake, however. For one thing, we may generally be justified in thinking that an event or state *A* triggers another event *D*, even if the influence of *A* on *D* is mediated by distinct states or events *B* and *C*. My pressing a key on a keyboard clearly, in a sense, ‘triggers’ the appearance of a letter on my computer screen, even though this influence is mediated by any number of internal states and events of my laptop.⁹ For another, the performance of some concrete bodily behaviour may be *guaranteed* long before the most causally proximal psychological state obtains; in such a case it would seem obtuse to insist that the latter state is the one

⁹ I thank both Will Davies and an anonymous reviewer for *Synthese* for bringing the significance of this point to my attention, and the same anonymous reviewer for this example.

that counts as *initiating* the behaviour in question. At the very least, we should admit that the behaviour has been triggered at the moment that its execution is guaranteed.

This suggests an alternative criterion; a psychological state s counts as initiating a bodily behaviour b just in case (1) s is rightly part of the causal explanation of b and (2) the persistence of s guarantees that b is attempted. This criterion intuitively distinguishes between the direct and supportive roles in action identified above. Though a visual representation of a ladder may be part of the causal chain that leads me to climb it, the persistence of such a representation in no way guarantees that I will try to climb it. A proximal intention to climb the ladder, however, satisfies both criteria.¹⁰

Though I admit that this criterion could benefit from more analysis, I think it provides enough of a guide for our purposes here, and so will avoid leading our discussion too far afield by analysing it further. In what follows, I simply adopt it.

2.2 Guidance

Next, I will clarify the notion of action guidance. Motivational mental states are generally thought to support the unfolding course of an action through to completion, by specifying both an agent's goal, and how it will be arrived at—an *action plan* (Pacherie 2012: p. 96). More precisely, a motivational mental state counts as guiding a behaviour just in case it (partially) determines the overall bodily movement that the agent exhibits in service of the represented goal.

I say 'partially' determines simply because it is typically held that motivational mental states specify a coarse-grained action plan (i.e. to kick the ball) rather than a fine-grained one (i.e. to bend the knee to some precise degree, move the foot forward at just such a speed, and make contact with the ball at some specific angle) (see Wu 2011: p. 60). That is, motivational mental states do not specify particular, individual motor trajectories, but abstract plans satisfiable by many different concrete movements. Thus any particular concrete bodily action is likely to be strictly underdetermined by the motivational mental state that gave rise to it. This speaks to a well-known 'inverse' problem in neurophysiology and motor control theory, which is usually referred to as 'Bernstein's Problem' (see Whiting's 1984 collection).

Nevertheless, it is thought that these plans play *some* role in specifying the final concrete movement. One noteworthy feature of the predictive account of motivation that I will present and defend in Sect. 5 is that it appears that motivational mental states *do* specify particular trajectories rather than merely abstract plans. That is, on such a view, the state in fact *does* fully determine (i.e. generate or predict) the precise trajectory of the relevant action, at least in a sense. I explain why this is so, and why we should not see it as a bug but rather as a feature of the predictive theory, in Sect. 5.2.

¹⁰ It is worth noting that this criteria will probably turn out to be inadequate if a proximal intention can persist in the face of overwhelming evidence that satisfying it is impossible. If I can intend to climb a ladder while also knowing that I cannot do so, then presumably the persistence of this intention will not guarantee that I will try. I am far from convinced that proximal intentions can persist in such circumstances (though perhaps distal intentions can) and so will not attempt to meet this challenge here. I will note, however, that any causal theorist of action who admits this possibility must suggest an alternative criterion by which we are to judge whether a state counts as initiating action. It is far from obvious to me what this criterion could be.

2.3 Control

Finally, I will clarify my working understanding of control. I think Josh Shepherd (2014: p. 400) has it right when he argues that an agent exhibits control over some action-type to the degree that they exhibit *flexible repeatability* in their performance of such actions. The action must be repeatable, because merely possessing the brute causal power to do something does not imply that one has any significant degree of control over that ability (Shepherd 2014: p. 400). Since I have the ability to kick a ball a short distance in front of me, I (strictly speaking) have the ability to kick a ball a small distance in front of me such that it goes into a football net. Nevertheless, I lack the ability to kick a ball in any particular direction with any degree of repeatability, and so I lack the ability to reliably score penalties. Intuitively, what I lack is sufficient *control* over my kicking.

Moreover, repeatability in identical sets of circumstances is very easy to come by. Imagine I can repeatedly score a penalty when I am perpendicular to the goal, but not when I change my angle with respect to the goal by a few degrees. Once again, it appears that I do not, in such a scenario, exhibit any significant degree of control over my kicking behaviour. My ability to perform the action must be repeatable in circumstances that differ in “some theoretically interesting way” (Shepherd 2017: p. 266). We need not investigate here exactly what ‘theoretically interesting differences’ amount to. Suffice it to say that the greater the degree of repeatability of some action, and the wider the range of circumstances in which that repeatability is exhibited, the greater the degree of control an agent exhibits over that action-type.

Note that the mere fact that an agent has only performed an action once does not preclude the agent exhibiting control on this account. The relevant repeatability and flexibility may be specified with respect to a well-selected collection of counterfactual circumstances (Shepherd 2014: pp. 400–403)—i.e. an agent may be said to exhibit a reasonable degree of control over an action if they *would* be able to repeat performance of the action in a reasonable range of counterfactual circumstances.

In the sense that I will use it here, a motivational mental state’s control *function* is satisfied when the state guarantees a significant degree of flexible repeatability in the performance of the behaviour it initiates and guides. It is generally thought that this is achieved by the motivational mental state being somehow responsible for the online monitoring of the progress of the relevant bodily movement, and for correcting this trajectory where it deviates from the represented action plan (Pacherie 2012: p. 96). That is, control, *qua* function of a motivational mental state, is achieved by two distinct processes; (a) monitoring and (b) online adjustment. For ease of referral, call these two distinct kinds of processes *components* of the control function.

It is worth reflecting for a moment on these two components. Monitoring seems to be somehow perceptual in nature, while online adjustment is specifically behaviour-oriented. This means that control, *qua* function of motivational mental states, straddles these two cognitive domains. Recognising this division in the standard understanding of the control function will be critical in what follows, since in my view Wayne Wu has a good grasp on the perceptual component of control (monitoring) and Bence Nanay has a good, though implicit, explanation of the behavioural component. This will lead

me to unify these two disparate insights under the rubric of predictive processing and active inference, in order to provide a full account of the control function.

3 Why existing theories are not enough

In this section I shall survey two existing theories of motivational mental states, the first due to Bence Nanay (2013) and the second due to Wayne Wu (2011, 2016). While I note that each provide helpful insight into the control function that other theories ought to preserve, I ultimately argue that neither satisfy the minimal criteria articulated above. This justifies the introduction of an alternative theory that broadly preserves each authors' insight while also filling the gaps. I shall introduce and defend such a theory in Sects. 4 and 5. Thus, this survey has two functions; to retain what insight we can into how we should go about satisfying our minimal explanatory demands, and to highlight the need for a novel theory if we are actually to do so.

3.1 Nanay on pragmatic representations

In his 2013, Bence Nanay puts forward a theory of 'Pragmatic Representations', which he characterises as "the representational component of the immediate mental antecedent of actions [IMAA]" (Nanay 2013: p. 17). By IMAA, Nanay refers to what we have so far called a motivational mental state; the IMAA is whatever mental state/states "trigger (or guide, or accompany actions)" (15). Further, Nanay suggests that the IMAA is divided into two components; "one that represents the world (or the immediate goal of the action...) in a certain way, and one that moves us to act" (16). He explicitly limits himself to providing a theory only of the *first* (cognitive) component, and not the second (conative) component. Hence, he explicitly distances himself from providing a theory (in our terminology) of the initiation function of motivational mental states. While this renders his theory necessarily partial by our lights, it is worth reflecting on what his account may provide to our understanding of the guidance and control functions of such states.

So what, according to Nanay, are pragmatic representations and what do they represent? He writes,

...pragmatic representations attribute properties [to external objects], the representation of which is necessary for the performance of an action...I call these properties... 'action-properties'. (Nanay 2013: p. 39)¹¹

More specifically, action-properties are properties of objects that we are trying to act on (e.g. by picking them up), that might include such things as location, size, weight, which all must be represented in an action-relevant format (i.e. weight must be represented in a format that allows me to work out what the appropriate force to exert is when attempting to pick it up). These are generally relational properties that

¹¹ It is part of Nanay's wider project to argue that Pragmatic Representations are perceptual states. Since nothing I have to say about his account depends on the stance one takes on this particular issue, I shall gloss over it here.

are nevertheless *attributed* to the external object; pragmatic representations represent the location of objects of action *relative to our location*, their size *relative to our grip size*, and their weight *relative to our strength* (2013: pp. 39–40).

So pragmatic representations are, according to Nanay, states that attribute all ego-centric, relational properties of objects necessary for us to act upon them, in a format appropriate for guiding the action in question. They are not, however, *sufficient* for action. To suffice for the generation action, pragmatic representations must be combined with some, independent and non-representational¹² (2013: p. 71) conative component of the IMAA/motivational mental state, which Nanay does not elaborate upon. Since the persistence of pragmatic representations do not guarantee that action will occur they do not count as initiating action by our lights, and it is clear that Nanay would agree. So the question that remains is what, if any, insight this theory can provide into the guidance and control functions of motivational mental states.

3.1.1 Nanay on guidance and control

I think it is fair to say that Nanay never explicitly lays out his views on the contribution of pragmatic representations to action guidance or control (at least not in these terms). But he nevertheless suggests that they are critical for at least the former function. He writes that,

Pragmatic representations can be correct or incorrect—any [action properties] can be correctly or incorrectly represented. If they are correct, they are *more likely to guide our actions well*; if they are incorrect, they are *more likely to guide our actions badly*. (Nanay 2013: p. 18)

Given his lack of specificity, it is reasonable to suggest that he does not have a well-developed theory of how pragmatic representations play (or contribute to) the guidance function. This is not necessarily a criticism of Nanay's project *per se*—he is, after all, striving for a theory of the nature of the representational component of the IMAA (read: motivational mental states), which does not explicitly require explicitly explaining how such states play all of their characteristic functions. Rather, it requires providing enough detail about the nature of pragmatic representations such that, *in principle*, one is able to explain whatever functions of motivational mental states the cognitive (rather than conative) component of motivational mental states is responsible for. Most plausibly, these are the guidance and control functions.

So, what does Nanay's analysis of pragmatic representations offer us for explaining action guidance? It is clear enough that the representation of all necessary action-properties is not itself an action-plan; knowing where my mug is, its size, and its weight does not constitute a plan to pick it up. Nevertheless, this content of a pragmatic

¹² As we will see, my preferred theory implies that Nanay is wrong to think of the conative component of motivational mental states as either independent of the 'cognitive' component or non-representational. Not only does the state responsible for action-initiation on my view represent action-properties, but further representational content (of the predicted proprioceptive consequences of movement) is *required* for it to perform this function. Thus he is also probably wrong to limit his examples of action-properties to those where properties are attributed to external objects (rather than to one's own body). I shall not dwell on such disagreements here.

representation is *necessary* for calculating an effective concrete motor trajectory, and it may also help to narrow the range of possible options. Certain action-types will be precluded by an object's represented location, distance, size, or weight. Moreover, the way in which these action-properties are represented may well determine, or at least suggest, a coarse-grained action plan; representing an object's weight as it relates to my grip strength suggests a different goal of behaviour than does representing an object's weight as it relates to my kicking power.

As I discuss in the following section on Wu, I do not think that this sort of suggestion, wherein motivational mental states constrain the space of possible motor trajectories so that a single concrete trajectory can be tractably calculated, is the best explanation available of their role in action-guidance. Nevertheless it is right, for now, to note that such a view is at least a *plausible candidate* explanation of the guidance function.

What about the control function? Here I think that Nanay has half the story (the other half will be provided by Wu, and the two will be plausibly combined in my own preferred account).

Recall that in describing the control function we identified two sub-functions; the *monitoring* function and the *online adjustment* function. I do not think that Nanay's pragmatic representations can tell us much about the former; it is not clear how these states can monitor actions in progress and detect deviations from an action plan, but they can provide some useful insight when it comes to the latter function.

Performing the online adjustment requires that adjustments are made to the action plan in response to perceived deviations. Assuming that pragmatic representations can play some role in determining what this action plan is (as described above), it follows that changes to the action-properties attributed to objects by a particular pragmatic representation could, in principle, be expected to alter the action plan.

This would work as follows. I form an action plan for picking up the mug on my desk in front of me. As I reach for it, my colleague nudges it with their elbow, knocking it several inches to the left. By whatever mechanism (see the section on Wu below for my view) I detect this change. Consequently, my pragmatic representation comes to attribute a different egocentric location, and perhaps distance, to the mug. This changes the possible range of motions that will achieve my goal of picking up the mug, and my action-plan is updated accordingly. Thus, I alter my arm's trajectory and (in the good case) nevertheless successfully pick up the mug.

As we will see, this story will receive a slightly different gloss when I propose my own preferred view in Sect. 5. Nevertheless, this Nanay-inspired view on the mechanism of online adjustment deserves significant credit for inspiring the proposal. For now, however, we will leave Nanay's partial theory of motivation behind, and consider an alternative proposal, due to Wayne Wu.

3.2 Wu on many-many problems

In his 2011, Wayne Wu proposes an account of how motivational mental states play their causal role in generating action. This is part of a broader project of identifying necessary and sufficient conditions for intentional bodily agency and control. The

following is a paraphrase of Wu's view of the mechanism by which motivational states perform their role in generating action.

An agent's (A) motivational state, M , plays a causal role in the generation of an action, φ , by structurally causing the selection required in implementing a solution to the non-deliberative Many-Many Problem appropriate to A φ -ing in the given context. (2011: p. 68)

Naturally, this requires some unpacking. Let us start with the non-deliberative Many-Many problem. This is a problem necessarily faced by any creature that is able to exhibit bodily agency. In any given situation, an agent will confront innumerable perceptual inputs, and a similarly staggering number of possible bodily behaviours. To act, an agent must, on the one hand, select a target from amongst the many identified by their perceptual inputs (that is, they must *attend* to some aspect of their perceptual field). On the other hand, agents must also select one from the many possible bodily behaviours they could perform (Wu 2011: p. 56).

A solution to the Many-Many problem, then, is a selection of both a target for action and a behaviour to perform on that target (roughly, e.g., paying attention to a hammer, and choosing to pick it up). The deliberative variant of this problem arises in the context of planning and practical reasoning/reflection. The non-deliberative problem (the important one for our purposes), on the other hand, concerns *implementation*. An easy way to think about the distinction is as connected to the distinction between distal and proximal intentions discussed in Sect. 1. A solution to the deliberative problem involves, roughly, the formation of a distal intention, whereas a solution to the non-deliberative problem involves a proximal intention (or relevantly similar motivational mental state) playing its characteristic functional roles in action. In Wu's terms the non-deliberative problem "is solved in part by the agent's exercising perceptual and motor capacities" (2011: p. 56). That is, these perceptual and motor capacities are supposed to underpin the initiation, guidance, and control of action. Though I agree that Wu has a good story to tell about the latter function, I shall raise problems below for his theory's explanatory adequacy regarding the first two.

We also need to better understand the notion of a structural (or structuring) cause. Wu's understanding is a slight expansion on that of Dretske (1993), who coined the term—for Wu, a structuring cause is an event or standing state "that produces certain enabling conditions whereby one event... can cause another..." (2011: p. 58). For Wu, motivational mental states are standing states that are also "structuring causes enabling specific links between attention and movement that amount to a solution to the non-deliberative Many-Many Problem" (2011: p. 58). That is, they *enable* the selection of perceptual inputs and behavioural outputs appropriate to the execution of the action that they specify, given the context in which the agent finds themselves (roughly, e.g., paying selective attention to a hammer and selecting an appropriate action; e.g., picking it up). This description of motivational mental states' role in action will be critical in what follows; in particular I think it leaves Wu's theory without a proper account of such states' initiation function. That is, I believe it is unclear why the persistence of intentions should, on Wu's theory, guarantee that the behaviour they represent will occur

3.2.1 Wu on control

In his 2016, Wu argues that causing the task-relevant selection of particular objects relevant to the performance of an action is a significant part of the way in which motivational mental states play their control function. I basically agree with Wu on this point,¹³ and will return to the main idea shortly. The first point worth noting however, is that Wu's theory of motivation involves positing that motivational mental states cause the deployment of attention towards task-relevant objects. The upshot, if we agree with the basic thrust of his 2016 paper, is that his 2011 paper can be read charitably as having offered an explanation of how motivational mental states perform their control function.

Wu (2016) begins by noting that the causal theory of action has classically been assailed by problems of causal deviance. Roughly, these are cases where motivational mental states cause bodily behaviour, but in the wrong way for that behaviour to count as intentional action. As Wu puts it, in such cases “behavior and its effects, even if intended, seem to happen accidentally” (2016: p. 103). For example, imagine a scenario where I am driving, and form an intention to kill my father. Realising that I have formed such an intention, I become seriously unnerved, so much so that I drive dangerously fast. This results in me knocking down and killing a pedestrian, who turns out to be my father (this sort of example is due to Roderick Chisholm (1966)). Plausibly my intention caused me to kill my father, but seemingly not in the right way. It does not seem I exercised control over the behaviour that ended up with the death of my father—this sequence of events certainly would not mean that I exhibit flexible repeatability with respect to the killing of my father. So this is not a case where I *acted* so as to kill my father, because my intention did not cause the killing behaviour ‘in the right way’. More generally, cases of deviance are ones where “some control-undermining state or event occurs between the agent's reason states [read: motivational mental states] and an event produced by that agent.” (Schlosser 2007: p. 188).

The upshot is that if we can identify a mechanism that operates only in cases of non-deviantly caused action (and plausibly explains such actions' lack of deviance), then we have identified a good candidate for a contributor to action control.

Wu claims, roughly, that cases where control is undermined are ones where the motivational mental state does not cause the deployment of attention to task-relevant elements of the environment. The consequence is that a primary source of information relevant to the monitoring and correction of the agent's movements is not available—effectively monitoring and correcting bodily movements depends on having information about the intended external targets of those movements. This position seems plausible when we reflect on paradigmatic cases of causal deviance. Where I am simply unnerved by my intention to kill my father, the intention does not cause me to attend to those aspects of the environment relevant to monitoring my movements' progress towards the goal-state, nor do I engage in any such monitoring relevant to achieving the outcome. Indeed, it is plausibly due to my *lack* of attention to the ‘target’

¹³ I disagree with him on a few details regarding the most plausible neurobiological mechanisms of attention. This disagreement is insignificant for the purposes of this paper so I will not dwell on it.

of my intention (but rather, perhaps, to my own nervousness) that I end up hitting my father. More generally, my failure to attend to task-relevant features of the environment will ensure that in a suitably selected collection of counterfactual circumstances, I will *fail* to kill my father. My intention does not perform its role of redirecting attention appropriately, thus I fail to exhibit flexible repeatability in my killing of my father and so, definitionally as well as intuitively, my control over my behaviour has been undermined.

These considerations count as some evidence that intentions (and other motivational mental states) may contribute to action control by causing attentional deployment towards the target of the action (as represented in the intention's content). Paradigmatic cases where we fail to exhibit control (or at least fail to exhibit *much* control) are ones where our intentions do not cause appropriate task-related attentional deployment.

Moreover, there is independent empirical evidence for this claim. For instance, Mrazek and colleagues (2013) found that, compared to controls, participants who partook in 2 weeks of 'focused-attention meditation' training, significantly improved their performance in a verbal reasoning and working memory task, and reported fewer and less severe instances of mind-wandering while they were performing the task. The suggestion here is that honing the process of intention-caused attentional deployment improved agents' control over the actions necessary to complete certain tasks. While these results are instructive, we must of course note the complication that the actions primarily involved in completing the aforementioned tasks were mental rather than bodily. More embodied examples (where participants are instructed to respond physically to, or withhold response from, certain cues appearing on a screen), also reveal a similar effect. Disengagement of attention from task-relevant aspects of the external environment, as might be expected, leads to impaired task performance (McVay and Kane 2009; Smallwood et al. 2008). Assuming that impaired task performance is a good measure of a lack of agential control over the task (which seems plausible, and is done elsewhere, see Shepherd (2017)), these results indicate that attention makes a crucial contribution to action control, making its selective deployment a plausible mechanism by which motivational mental states play this functional role.

3.2.2 Wu on guidance and initiation

Such is Wu's suggestion when it comes to motivational mental states' control function. How does he account for their role in action guidance? According to Wu, motivational mental states aid the computation of a suitable concrete movement by significantly constraining the initially enormous space of motor possibilities from which that concrete movement is to be selected. In other words, motivational mental states, when they function correctly, impose prior constraints in order to resolve, effectively, Bernstein's problem; 'how should I move next, given the huge number of options'. This, in turn, makes it tractable for a distinct subsystem (2011: p. 60) to compute, for the constrained set of concrete movements, those that minimise the value of some cost function, which typically amounts to computing which movement minimises the value of some (or several) relevant motor parameter(s), such as distance travelled to target, or jerk (2011: p. 60).

This work of constraining the computational space is purportedly achieved thanks to the representational content of the intention—intending to perform a movement *with your right arm*, for instance, greatly reduces the number of possibilities from amongst which movement must be specified. That is, selection is limited to contraction or relaxation of the muscles in the right arm, and to those muscular activations in the right arm consistent with the general character of the arm movement intended (for instance, the contraction of the bicep in the case of a ‘curling’ motion). This constraining role can be seen as the abstract specification of an action plan, which is transformed into a concrete movement by a further computational process. This picture is, I grant, a *prima facie* plausible explanation of motivational mental states’ guiding function.

On the motor side of the Many-Many problem (as opposed to the perceptual side), Wu describes motivational mental states’ role *only* in terms of constraining the set of motor possibilities, as described above. That said, it sometimes seems like Wu envisions the outcome of cost-function computation (ideally, the specification of a single concrete movement that minimises the value of the cost-function) as the initiation of action, though the details are never made clear. Wu writes,

Once the lowest cost movement is identified, the agent’s body moves and does so in accordance with his intention (2011: p. 60)

I think that any plausible interpretation of what Wu is suggesting here leaves him without an explanation of action initiation. I shall explain this thought in more detail in the following section, but it is worth noting at this point that identifying how Wu’s story is supposed to account for action initiation is difficult at best. Given Wu’s goals in his 2011 paper, this is an understandable omission. But it does, I shall argue, make it difficult to think of him as providing a complete theory of motivation. What Wu has provided on the matter is, on the surface at least, insufficient to scratch that particular itch.

3.2.3 Wu: the problem of initiation

Problems emerge when considering Wu’s theory’s ability to account for motivational mental states’ initiating function. The problem is that after motivational mental states have constrained the set of motor possibilities in the manner Wu describes, all subsequent work is attributed to the computation of cost functions.

Either Wu must accept that one of these two stages explain action-initiation, or he must immediately accept that there is no element of his theory that explicitly explains initiation.¹⁴ Since the second disjunct leaves his theory critically incomplete, the most promising option is to accept the first. The problem is that this does not solve his problem.

¹⁴ This is, of course, not to deny that his theory could be augmented by positing *another* state that explicitly explains initiation. It would take us too far afield to consider what kind of amendments might save Wu’s account here, but note that it would have to combine with the existing posits of his theory (or work independently) to *explain* initiation, i.e. *why* the persistence of this new state guaranteed action. It would not be enough to simply *stipulate* that a different state is responsible for initiation. I point this out not at all to suggest that such a task is impossible, but to highlight the amount of substantive theoretical work that would, if I am right, need to go into responding to this objection. I thank an anonymous reviewer for *Synthese* for giving me the opportunity to clarify the intended force of my objection here.

Recall that we earlier agreed that explaining action-initiation amounted to explaining why the persistence of a given motivational mental state (or a sub-component) guaranteed that a behaviour would occur. To answer this challenge, it seems as if Wu could begin by saying one of two things; either (1) the persistence of the end-state of cost-function computation guarantees action (this seems to be a reasonable gloss of what he says above) or (2) the persistence of whatever state triggers the process of cost-function computation (i.e. the end-state of the process of constraining motor possibilities) guarantees action.

The issue is that both of these options are unhelpful, for the same simple reason; there is no reason why *either* should guarantee action. Given everything Wu says, there is nothing about the output of either cost-function computation or the process of constraining motor possibilities that would guarantee that action occurs (consider—why should eliminating a number of action possibilities, or selecting a particular concrete motor trajectory, guarantee that a concrete movement actually occurs?). Naturally, Wu could simply *stipulate* that one or the other guarantee the performance of behaviour, but this would be entirely *ad hoc*, and, in any case, would not serve to *explain* action-initiation. Such a stipulation, even if true, would give us no insight into what is special about the state in question; what *differentiates* it from the wide range of mental states whose persistence *does not* guarantee behaviour.

I believe that the predictive theory I offer below *can* differentiate such mental states, but will briefly turn to another criticism of Wu, before beginning to spell it out.

3.2.4 Wu: the problem of guidance

So much for action-initiation. How does Wu's account of action-guidance fare?

One thing to note is that the notion that mental states have a constraining role in action guidance could, from one perspective, be challenged by arguing that it is trying to solve a problem that is not there. As we will see later, Bernstein's Problem (which Wu's story is implicitly trying to resolve) does not emerge if action is prescribed by the predictions of a forward model. Interestingly, in the neurosciences, some think that Bernstein's Problem has been effectively dissolved with the advent of the Equilibrium Point Hypothesis, the Passive Movement Paradigm, and other ideas (see Latash 2010; Mohan and Morasso 2011). To simplify enormously, these approaches avoid the need for any constraints, by presenting motor control as a problem of specifying the *consequences* of an action. We will return to this general point later when considering active inference in predictive processing.¹⁵ For now, however, we will not pursue this particular line of thought any further in objection to Wu.

While Wu undoubtedly has *some* account of the mechanism underpinning the guidance function, it is worth noting that some of its key elements are increasingly disputed. In particular, the claim that motivational mental states' guidance role is achieved by enabling the efficient computation of cost functions should be treated with suspicion. Many working roboticists (Feldman 2009 and Mohan and Morasso 2011), as well as neuroscientists and philosophers (Friston 2011; Clark 2016: p. 132) have branded

¹⁵ I thank an anonymous reviewer for *Synthese* for carefully explicating this helpful way of viewing the disagreement between Wu's kind of view and that put forward by many advocates of active inference in Predictive Processing.

explicit cost-function-based solutions to the problem of efficient motor selection inflexible and biologically unrealistic. As Mohan and Marasso put the issue,

...such engineering paradigms were designed for high bandwidth, inflexible, consistent systems with precision sensors. The difficulty lies in adapting these models to the typical biological situation, characterized by low bandwidth, high transmission delays, variable/flexible behavior, noisy sensors, and actuators. (2011a: p. 3)

Consequently, we should be suspicious of whether, despite their explanatory success, models involving the explicit and resource-intensive process of cost-function calculation are likely to reflect what is actually going on in complex biological systems such as ourselves. There is, for instance considerable doubt in the literature regarding whether the impressive formal solutions to many *prima facie* problems faced by these models can actually be efficiently implemented through distributed neural networks¹⁶ (Todorov 2006; Mohan and Morasso 2011: p. 3). One upshot of these concerns ought to be suspicion of the claim that motivational mental states play their guidance function by enabling the activity of these biologically dubious constructs.

Let us say, however, that it could be demonstrated that these models were, as far as we can tell, biologically implementable. There is still a further difficulty. Wu's proposed solution presumes that the representational content of intentions is sufficiently detailed so as to narrow the space of motor possibilities to a degree that makes the computation of a cost function *tractable* for each possible residual motor behaviour. Maybe this is so, but there are good reasons to doubt it. For instance, the initial set of possible concrete motor behaviours numbers in the region of 2^{600} (Wolpert and Ghahramani 2000). Solving a cost-function for even a tiny portion of that space is phenomenally computationally expensive. Consequently, intentions would need to be having an *enormous* constraining effect on this behavioural space for our actions to proceed at the rate at which they typically do. It is not clear, given the characteristically coarse-grain at which we think of the content of our intentions (and other motivational mental states), that this is a safe assumption.

There is no knock-down argument against Wu's proposal here (or even in the vicinity), absent further empirical study of action-generation. But it is possible to make sensible theoretical decisions in the meantime. In particular, an alternative theory that did not posit the computation of cost functions as a key mechanism would have the potential to appear significantly more biologically realistic, and less of a hostage to empirical and computational fortune. By extension, a theory of motivational mental states' guidance function that does not place them at the centre of an overly computationally complex and biologically unrealistic process should, I think, currently be preferred over one that does, all else being equal.

In Sect. 5, I argue that my predictive theory of motivation satisfies this requirement, as well as offering an account of motivational mental states' role in action-initiation and preserving Wu and Nanay's insights on the topic of control. Now, however, I turn

¹⁶ The reasonable assumption being that such models are our best guess at how the brain 'computes', to the extent that it does at all.

to providing the background information required to understand the theory posited in Sect. 5.

4 Predictive processing and active inference

In this section, I outline the key features of the predictive processing (PP) research paradigm in Cognitive Neuroscience. I do not attempt a systematic defence of the framework. I take that it has been well enough articulated, philosophically defended (see Clark 2013a, b, 2016; Hohwy 2013), and empirically supported (Talsma 2015; Barrett and Simmons 2015; Seth 2013; Adams et al. 2013) to warrant at least being taken seriously. Instead, I explain the framework here, and make use of it in the articulation of the predictive theory of motivation in Sect. 5.

The central concept that the theory of PP brings to the table is **prediction error**. This, roughly, denotes a signal hypothesised to be produced when the brain's predictions about the sensory signal (broadly construed to include exteroceptive, interoceptive, and, crucial for our purposes later on, proprioceptive signals) fail to match the actual sensory signal. The overall structuring principle of the human mind, according to PP, is the minimisation, over the long-term, of prediction error signals.

Let us call the brain's best current guess regarding the source of the incoming sensory signal (i.e. its causal origins) the **world model**. On the basis of the **world model**, the brain produces a **generative model**; “a flow of virtual or mock sensory signals that predicts the...sensory signals generated by external¹⁷ causes” (Gładziejewski 2016: p. 562). The generative model is hierarchical in nature; the highest level of the generative model is generally thought to encode relatively abstract predictions about the expected state of the world, which unfold as the hierarchy descends into relatively concrete, precise predictions about the expected character and intensity of particular sensory signals. For instance, a high-level prediction that an elephant is very close in front of you might unfold into particular visual (e.g. grey, textured), auditory (e.g. trumpeting, stomping), and olfactory (e.g. strong, muddy) predictions at the lower levels of the hierarchy.

Those aspects of the *incoming* sensory signal that are successfully predicted by the generative model are dismissed—that is, they have no further effect on any aspect of the neural economy of which they were a part. They have been, to use a term of art, successfully ‘explained away’. Those aspects that are *not* explained away, on the other hand, generate a further upward-flowing signal. This, explained slightly more precisely this time, is the prediction error.

The function of prediction error signals is to drive mental processes that have the effect of minimising future prediction error. At any given time,

...the cognitive system attempts to settle on a “hypothesis” about their [the sensory signals’] causal origins, namely that which has the highest posterior probability (i.e. the probability of being true in light of the data) among

¹⁷ In fact, this qualifier does not always apply to the predictions of the generative model; it is crucial to the proposal I will make later in the paper that the sensory signals it is predicting do not always have external (at least in the sense of ‘outside the boundaries of skin and skull’) causes.

alternatives, given its likelihood...and prior probability... (Gładziejewski 2016: pp. 561–562)

The result of this process of ‘settling’ on a hypothesis (that is adopting a hypothesis that generates minimal prediction error), according to PP, corresponds to what we perceive, believe, do, and so on. The reason why the process of settling can produce all of these features of human mentality is that there are, PP advocates are keen to point out (Clark 2013a, b; Friston 2011), *two ways* for embodied, active neural systems such as ourselves to minimise prediction error. The first way is for generated prediction error to drive updates to the world model (and hence the generative model). That is, the brain engages in a process of (roughly) Bayesian inference to the best explanation of the prediction error signal, and adopts a new world model that (when all is going well) produces a generative model that results in less prediction error than before. Call this (Bayesian Belief updating or) **passive inference**.

The second way operates in the other direction; instead of driving internal changes to the world model, generated prediction error drives changes to the source of the sensory signal (i.e. the world), in order to bring it more in line with the predictions of the generative model. That is, generated prediction error can cause us to *act* on the world in a way that alters the sensory signal so as to bring it more in line with what is predicted. Call this (World updating or) **active inference** (Clark 2013a; Friston 2010).

We should not, however, expect these processes to entirely eliminate prediction error, even if we suppose the world model is a perfect representation of the distal causal structure of the world the agent is currently occupying. This is because the typical sensory signal (generated as it is by imperfect sensory transducers) is *noisy*—that is, the signal will typically partially misrepresent the state of the world.

This poses a problem. Given that we typically get around in the world successfully, it seems like our brains are, for the most part (though certainly not always) able to distinguish between prediction error that it *needs to do something about* and that which it does not. What might allow the brain to distinguish between prediction error born of mismatch between the world model and reality, and that born of mismatch between the world model and sensory noise?

The answer involves postulating further predictive processes. As well as generating predictions regarding incoming sensory signals, the generative model also produces predictions regarding the expected quality of the incoming sensory signal (i.e. how noisy/clean it is) as well as its own predictions/mock sensory signal. This is known as the **expected precision** of the signal. Various contextual features, as well as past experience regarding certain kinds of sensory signals, are used to calculate a particular signal’s expected precision (Clark 2013a, b, 2016; Gładziejewski 2016: p. 562). The proposed details need not concern us here. The crucial point is that both active and passive inference are significantly more likely to be driven by prediction errors driven by *high-precision* sensory signals or predictions of the generative model. Prediction errors resulting from *low-precision* signals/predictions have significantly less power to drive update or action.

If a sensory signal has low expected precision, as might occur, for instance, in the case of visual signals on a very foggy day, it will not drive processes of perceptual or active inference. In the PP worldview, the process of modulating the expected

precision of the sensory signal (and thus the impact of related prediction errors on the brain's inferential processes) in this way, corresponds (at least roughly) to the notion of *attention* as it is understood in psychology and neuroscience more widely.

Having explained the basics of predictive processing, in the next section I present the predictive account of motivation, and suggest that it is to be preferred to both Wu and Nanay's accounts, according to the standards previously identified.

5 The predictive theory of motivation

As we have seen, according to the predictive processing framework, action (or active inference) is just another manner by which the brain-body system minimises overall prediction error. One central upshot is that action is the outcome of basically the same processes as perception, or belief. But it is important not to jump from this observation to the claim that there is no significant difference between motivational mental states and motivationally inert ones. As Shea notes (2013) we still have to account for why certain prediction errors are minimised passively (i.e. by revising the predictions of the generative model), and why others are minimised actively (i.e. via environment-altering action). To put the point another way, the mere possibility of action is insufficient; we need the predictive processing framework to give us a plausible story of how action *actually* comes about, not merely to gesture in the general direction of why the occurrence of action is *compatible* with the overall story being told.

I take it that the complete story is yet to be told. But friends of the predictive processing story can certainly point out one kind of situation where the model predicts that agents will act, rather than revise their generative model. This is one in which one or more of the highest-level predictions of the generative model are assigned arbitrarily high precision in comparison to the incoming sense data with which they might conflict (Klein 2018). Any prediction errors generated by conflict between such predictions and incoming sense data *must be* resolved actively, because the degree of precision assigned to the prediction of the generative model renders it *effectively un-revisable*. Intuitively speaking, we can think of this as a situation in which the prior probability of the prediction is judged so high that no realistic amount of conflicting sense data can drive the highest levels of the cognitive hierarchy to revise it. In such a situation, passive inference is not an option for minimising prediction error. Since the overarching goal of the system to minimise prediction error is still in place, the only remaining possibility is active inference (i.e. action) to make the relevant prediction true.

The central claim of the predictive processing account of motivation I will be putting forward here is that some kind of dependence relation¹⁸ holds between these effectively un-revisable predictions of the generative model and our motivational mental states. Hence we have moved from the mere possibility of action, to actual action; motivational mental states actually cause action, and are hence distinguished from motivationally

¹⁸ I refrain from speculation regarding the exact nature of this dependence relation. It could turn out to be identity, determination, ground, supervenience, or perhaps something else. Nothing that I have to say in what follows requires me to take a position on this question.

inert states, because they generate prediction errors that can *only* be minimised through active inference. I also hold that this theory of motivational mental states can more successfully account for the functional roles of motivation than Wu's story.

In order to obtain this result, however, I must add one more posit to the theory. We are currently working on the supposition that motivational mental states depend closely on the activity of the highest-level predictions of the generative model. We must say one more thing about these predictions; they decompose (at lower levels of the hierarchy) into at least two kinds of prediction; *high-precision* proprioceptive predictions and (perhaps surprisingly) *low-precision* exteroceptive predictions. The proprioceptive predictions can be thought of, intuitively, as predictions of the expected proprioceptive consequences of actually acting in the present situation so as to satisfy the content of the motivational mental state. So, for instance, if I presently intend to pick up my coffee cup, the associated high-level prediction that I *will* pick up my coffee cup gives rise (at a lower level of the cognitive hierarchy) to a high-precision prediction of the proprioceptive consequences of actually moving my arm in the relevant direction and performing a grasping motion. The exteroceptive predictions can be thought of, intuitively, as predictions of the location and (action-relevant) qualities of task-relevant features of the external environment. So, continuing with the same example, the associated high-level prediction gives rise (at a lower-level of the cognitive hierarchy) to a low-precision prediction of the ego-centric location and circumference of the coffee cup (amongst other features).

To summarise; with the following two posits, and the background assumption that the cognitive system functions so as to minimise overall prediction error, I believe I can offer a satisfying account of the functional roles of motivational mental states.

1. A dependence relation holds between motivational mental states and high-level, effectively un-revisable predictions of the generative model.
2. Effectively un-revisable predictions of the generative model decompose at lower levels of the hierarchy into (a) *high-precision predictions of the proprioceptive consequences* of acting so as to satisfy the motivational mental state in the current situation and (b) *low-precision predictions of the action-properties* of external objects and environment features (e.g. object-size relative to grip strength, ego-centric location and distance, etc).

I will now detail the proposed explanation of each of the central functional roles of motivation in turn.

5.1 Initiation

We saw in Sect. 3 that Wu and Nanay's accounts failed to offer any satisfactory account of the initiation function of motivational mental states. According to the predictive theory, initiating an action is the consequence of the stable presence of an effectively un-revisable prediction that the change to the world associated with the action will actually occur. For instance, if I intend to pick up my coffee cup, then the initiation of my picking up the coffee cup is an unavoidable consequence of the persistence of the generative model's effectively un-revisable prediction that the cup is (or will shortly be) in my hand.

On the predictive theory motivational mental states determine, ground, or are identical to an effectively un-revisable prediction that *guarantees* active inference (i.e. action). Because the prediction is effectively un-revisable, and because the brain consistently functions so as to minimise overall prediction error, in such a situation bodily movement *must* occur (naturally, the bodily movement in question may nevertheless fail to actually satisfy the agent's goal). To put it another way, the stable presence of a motivational mental state is, in this context, sufficient to ensure that action is attempted, since there is no other way for the cognitive system to minimise (as it must) prediction error generated by the following conflict; the prediction that the coffee cup is in my hand, conflicting with the incoming sensory information that it currently is not.

Naturally, the guarantee of action in this situation is not a matter of metaphysical necessity. It is, rather, something like a matter of *cognitive necessity*; given the general principles governing the operation of human cognition (according to the predictive processing view), the stable presence of an effectively un-revisable prediction is sufficient for a bodily movement to be initiated that will attempt to make the world such that it conforms to the predicted state of affairs.

5.2 Guidance

As we also saw in Sect. 3, Wu and Nanay's accounts offered an unsatisfactory account of motivational mental states' guidance function. The central problem was that both stories invoked a computational process of simplifying the minimisation of some complex cost function. This sort of process turned out to be biologically unrealistic, and computationally expensive, in a manner that made it look implausible as a hypothesis regarding the actual processes underpinning motivational mental states' guidance function.

According to the predictive theory, motivational mental states are closely tied to effectively un-revisable predictions of states of affairs that decompose into (amongst other things) high-precision predictions of the proprioceptive consequences of bringing those states of affairs about. As Clark writes,

...motor control is, in a certain sense, *subjunctive*. It involves predicting the non-actual proprioceptive trajectories that would ensue were we performing some desired action. (2016: p. 121)

This is how motivational mental states satisfy their guidance function, or so I shall suggest. Instead of first restricting the enormous set of possible motor trajectories, and then computing a cost-function over this set to select the least costly concrete behaviour, the predictive theory proposes that the motor trajectory is already specified directly by the associated proprioceptive prediction. In other words, predicting the proprioceptive consequences of bringing about a certain state of affairs already amounts to having selected an executable action plan. This is because distinct trajectories are also proprioceptively distinct. A prediction of the proprioceptive consequences of a particular movement specifies a unique trajectory for such a movement (i.e. the trajectory that will result in exactly the flow of proprioceptive sensory information that is predicted). As Friston writes,

...in active inference, these problems are resolved by prior beliefs about the trajectory (that may include minimal jerk) that uniquely determine the (intrinsic) consequences of (extrinsic) movements.... (2011: p. 496)

And as Clark boldly states,

It is easy...to specify whole paths or trajectories using prior beliefs about (you guessed it) paths and trajectories! (2016: p. 130)

None of this, it should be pointed out, reduces the *overall* computational complexity of selecting from the intimidatingly large set of possible motor trajectories (Clark 2016: p. 130; Friston 2011: p. 492). It simply pushes the problem back to the acquisition of an effective model, which is good at predicting *actionable* proprioceptive consequences of bringing about certain states of affairs. This is certainly not a trivial task.

This solution is, however, a more realistic prospect than Wu's (and, at least implicitly, Nanay's) in at least one important way. On Wu's proposal, the problem must be solved in real-time, on each occasion the agent is motivated to act. On the predictive theory, on the other hand, this problem is being constantly solved, on each occasion the priors of the generative model for a given situation are updated in response to prediction error. That is, the time given over to solving this problem of serious computational complexity is considerably greater on the predictive theory than it is on Wu's. This should lead us to think that it will be significantly more tractable. This, I suggest, is a significant advantage of the predictive theory, which is otherwise able to do all the work of Wu's account of guidance.

In fact, this story about guidance does considerably more than Wu's, in a way that one might initially find troubling. Recall that when defining the notion of action-guidance, I noted that the most common way of thinking supposed that motivational mental states did not specify entire trajectories, but rather abstract action plans, satisfiable by many different concrete movements. That is, the action-guiding role of motivational mental states does *not*, it is thought, involve solving the many-one mapping problem of motor control.

Clearly, on the predictive account, the proprioceptive predictions achieve precisely this. They *uniquely* specify a particular motor trajectory, by way of such a trajectory's proprioceptive consequences. Indeed, the predictive story explains by way of one process that which Wu is forced to explain by way of two; while he must appeal to both motivational mental states' restricting the class of concrete actions and a cost-function computation over the remaining options, both of these processes are taken care of by high-precision proprioceptive predictions on the predictive account.

Whether or not this is a problem, I think, largely depends on how one looks at it. If you think that one's explanation of motivational mental states' role in action-guidance should not, in principle, imply that they also solve the many-one mapping problem, then this will strike you as a problem. But it seems likely that this restricted view on the limits of the action-guidance role was a product of precisely the kind of view that Wu advocates; one on which it was unclear *how* a motivational mental state could have a sufficiently fine-grained content to be able to solve the many-one mapping problem. Since it is now less mysterious how this could come about, we should not insist from the start that the two-step solution is the right one. Thus we should not

accept the restricted role of motivational mental states in action-guidance that such a view suggests.

One may well think, of course, that motivational mental states still *do not* have such precise content even on the view I am advocating. Effectively un-revisable predictions of the generative model needn't themselves have these high-precision proprioceptive predictions as content. What *is* true is that (based on past experience of such situations) they cause a downward flow of increasingly precise predictions that 'unpack' the likely consequences of that prediction, including the specific expected proprioceptive signal. Whether you should see this as amounting to the motivational mental state itself (or at least the effectively un-revisable prediction to which it bears some kind of dependency relation) having the relevant proprioceptive content, or simply causing a lower-level state that does, is unclear. But whichever one opts for is, for our purposes here, inconsequential. The necessary work gets done either way.

5.3 Control

I endorse Wu and Nanay's insights on how motivational mental states play their action-controlling function. The key ideas were (a) that paradigmatic cases of action-control being undermined were ones in which there was a failure of attention (which suggests a failure of monitoring) and (b) that an action plan could be changed on the fly by altering the representation of egocentric action-properties as monitoring reveals changes to the external environment. In this final substantive section, I shall argue that the predictive theory retains these insights, and thus gives a full account of control, where Nanay and Wu were only able to tell half the story.

To start with monitoring, recall that on the predictive theory, effectively un-revisable high-level predictions of the generative model decompose at lower levels of the cognitive hierarchy into (amongst other things) low-precision exteroceptive predictions of the action-properties of external objects and environmental features. This, I suggest, implies a recalibration of attention towards the relevant features of the external environment, in a manner that allows sensory information to drive updates in the world model. The reason is that since the predictions are of low-precision, conflicting sensory information will be of comparatively high-precision, and thus overwhelmingly likely to be treated as data, not noise.

The central point here is that if the predictive theory is correct, then motivational mental states generate exteroceptive predictions about action-relevant features of the task environment that are especially prone to update in light of conflicting sensory information. Thus throughout the execution of the action, the agent's world model is especially susceptible to being updated in line with the evidence of their senses. To put it simply, motivational mental states perform their action-controlling function by opening up the generative model to being revised by the task-relevant features of the environment. This seems to be a predictive twist on Wu's suggestion that the deployment of attention towards task-relevant features of the environment allows the online monitoring and modification of bodily behaviour. Incoming sensory information is permitted to 'correct' mistaken predictions about an object's location or other action-relevant features (because these predictions are assigned low-precision), which

will feed into the details of the associated proprioceptive predictions. Moreover, the fact that the low-precision exteroceptive predictions concern action-relevant features of the present environment means that *those properties* rather than any others are the ones being ‘monitored’. Thus the predictive theory can retain Wu’s insights into action control.

Some readers might reasonably think that there is a tension between PP and the attentional account of the monitoring component of the control function.¹⁹ PP mandates that the system estimates the precision of current sensory signals as low, so that it can ‘believe its own priors’ and engage in active, rather than passive, inference. The comparatively high precision of the proprioceptive component of the motivational mental state, according to my view, is what makes motivational mental states effectively unrevisable. Since PP equates (or very closely links) precision estimation with attention, it seems as if my view is committed to the consequence that action requires attending away from current sensory signals. But this seems to be at odds with the claim that control is partly achieved by attending to incoming sensory signals.

This tension is merely apparent, because it misconstrues what must be attended away from, according to the PP view of motivation I have put forward. Recall the point I made above regarding precision estimation being a zero-sum game; for action to occur, a system capable of active inference must assign a low precision estimate to incoming *proprioceptive* sensory information (and my view entails that it will, since it posits the assignment of *high-precision* proprioceptive *predictions*—these are two sides of the same coin in the PP model). My hypothesis, however, involves a system assigning low precision to downward flowing *predictions* of *exteroceptive* sensory information (and thus high precision to the incoming flow of exteroceptive sensory information).

These two things are entirely compatible, and the possible confusion underscores the importance of distinguishing between different kinds of sensory attenuation necessary for action, according to the predictive view of motivation. In summary; the system must attenuate proprioceptive sensory information, but *enhance* exteroceptive sensory signals (which it achieves by attenuating *its own predictions* in this domain).

Note that this chimes with theories of ‘choking’ in skilled performance that posit the cause as being excessive self-focus (read, excessive attention to bodily/proprioceptive sensory information) (see, e.g., Beilock and Carr 2001). Though competing accounts of choking exist, primarily those that suggest it is a consequence exteroceptive attention being directed at task irrelevant features (see, e.g., Eysenck and Calvo 1992). It is worth pointing out that (a) self-focus theories dominate the current literature on choking and (b) that recent work attempting to bridge the dichotomy between these competing accounts accepts that excessive self-focus can be *a* cause of choking, even if it is not the *only* cause (Christensen et al. 2015: p. 288). This suggests strongly that the literature is converging on the view that some attenuation of proprioceptive sensory information and enhancement of the *task-relevant* exteroceptive information is required for successful action, just as the predictive theory suggests, though naturally they may not suffice for it.

¹⁹ I thank an anonymous reviewer for *Synthese* for bringing this worry to my attention.

Now we can move on to consider online adjustment. Recall that I argued that online adjustment, inspired by Nanay's view, could be thought of as a process of updating the action-properties that Nanay thought pragmatic representations attributed to external objects, and consequent updating of the action-plan. Though making no mention of pragmatic representations, this is effectively what the predictive theory suggests will occur as a result of the interaction of low-precision prediction of action-properties and (consequently) relatively high-precision sensory information pertaining to those same action-properties. Since the predictions are of low precision, prediction error generated where they meet incoming sensory information will, generally, be resolved by updating the world model. Thus, as sensory input reveals changes in the predicted egocentric action-properties of objects and features, subsequent predictions of the generative model will reflect these updated action-properties, in a constant process of feedback and update, driven from the 'bottom-up'. For example, as my hand moves forward, the direction and distance of my mug (relative to my hand) will change in a way that conflicts with the initial predictions of these properties, resulting in prediction error that drives relevant updates in my world model.

These updates to the world model will not be isolated from the rest of the predictive economy; they will be used to inform updates to the expected proprioceptive consequences of my grabbing the mug (or whatever other behaviour). This would work in the following way; as changes in action-properties are predicted by the generative model in response to sensory prediction error, these changes may or may not result in adjustments to the expected proprioceptive consequences of grabbing the mug. Roughly, if the changes to represented action-properties are consistent with the predicted proprioceptive consequences of movement (i.e. if the egocentric distance to the mug has shortened because of a *predicted* movement of my hand), then the update will result in no changes to the expected proprioceptive consequences of grabbing the mug (since grabbing the mug is happening in just the way was initially predicted, proprioceptively speaking). But if the newly represented action-properties are *inconsistent* with the predicted proprioceptive consequences of movement (i.e. if the egocentric distance to the mug has changed because someone else unexpectedly nudged either my hand or the mug), then the expected proprioceptive consequences of picking up the mug will change along with them (since grabbing the mug can no longer happen in the expected way, proprioceptively speaking). After all, the expected proprioceptive consequences of picking up a mug partially depends on the egocentric distance, direction, and so forth, of the mug.

All of this is once again assured because of the effective unrevisability of the high-level prediction that the mug will be picked up; while this state persists, minimising prediction error necessarily involves adjusting for unexpected changes in objects' and features' action-properties (i.e. such changes that are not the expected consequences of a particular motor trajectory, specified proprioceptively). This is because the only other option is to, as a result of the perturbation, cease to predict that the mug will be picked up. But while this prediction is effectively unrevisable, this is not an option at all.

To summarise, the predictive theory preserves Wu's and Nanay's insights into the cognitive bases of the control function of motivational mental states, and unifies them into a single model. A motivational mental state produces low-precision predictions

of an object or feature's action-properties (which is equivalent to saying it causes incoming sensory information about those action-properties to be treated as relatively high-precision). This is, on the underlying predictive processing model of cognition, the same as saying that it causes the deployment of attention towards task-relevant environmental qualities (as Wu says). This way of monitoring the environment will result in continuous update to the action-properties predicted by the generative model as the action unfolds, which will in turn drive (where necessary) online adjustment to the expected proprioceptive consequences of the motivational mental state.

6 Conclusion

I have argued that any account of how motivational mental states cause action must account for how they play three specific functional roles; initiation, guidance, and control. I have also proposed a theory of motivation drawn from the predictive processing paradigm (the predictive theory), and argued that it was able to explain how motivational mental states perform all three functional roles in detail. Thus I conclude that it significantly improves on extant work in action theory due to Wayne Wu and Bence Nanay.

It remains to be seen whether there might emerge accounts, including perhaps some revised versions of the work I have reviewed here, that are able to offer more complete or otherwise better explanations of the functional roles identified as centrally important here. It is also an open question whether an adequate theory of motivational mental states must explain more than just these three functional roles. That said, what I have proposed here is a significant step forward, and undoubtedly more grist for the active inference mill.

Acknowledgements For their help inspiring, guiding, and developing this paper, I would like to thank Lisa Bortolotti, Will Davies, Iain Law, Casey Elliott, the audience at the Predictive Processing: State of the Art (and Just Beyond) workshop held at the University of Edinburgh in November 2018, and two anonymous reviewers for *Synthese*.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure and Function*, 218(3), 611–643.
- Bach, K. (1978). A representational theory of action. *Philosophical Studies*, 34(4), 361–379.
- Bain, D. (2011). The imperative view of pain. *Journal of Consciousness Studies*, 18(9–10), 164–185.
- Barrett, L., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16, 419–429.
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, 130(4), 701–725.
- Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.

- Chisholm, R. (1966). Freedom and action. In K. Lehrer (Ed.), *Freedom and determinism*. New York: Random House.
- Christensen, W., Sutton, J., & McIlwain, D. (2015). Putting pressure on theories of choking: Towards an expanded perspective on breakdown in skilled performance. *Phenomenology and the Cognitive Sciences*, 14, 253–293.
- Clark, A. (2013a). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204.
- Clark, A. (2013b). Expecting the world: Perception, prediction, and the origins of human knowledge. *The Journal of Philosophy*, 110(9), 469–496.
- Clark, A. (2016). *Surfing uncertainty: Prediction, action and the embodied mind*. Oxford: OUP.
- Corns, J. (2014). The inadequacy of unitary characterizations of pain. *Philosophical Studies*, 169(3), 355–378.
- Dancy, J. (1993). *Moral reasons*. Hoboken: Blackwell.
- Dretske, F. (1993). Mental events as structuring causes of behavior. In J. Heil & A. R. Mele (Eds.), *Mental causation*. Oxford: OUP.
- Elliott, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion*, 30(2), 111–116.
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, 6(6), 409–434.
- Feldman, A. G. (2009). New insights into action–perception coupling. *Experimental Brain Research*, 194(1), 39–58.
- Frijda, N. H. (1986). *The emotions*. Cambridge: CUP.
- Friston, K. J. (2010). The free-energy principle: A unified brain theory. *Nature Reviews Neuroscience*, 11(2), 127–138.
- Friston, K. J. (2011). What is optimal about motor control? *Neuron*, 72(3), 488–498.
- Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese*, 193(2), 559–582.
- Hohwy, J. (2013). *The predictive mind*. Oxford: OUP.
- Klein, C. (2018). What do predictive coders want? *Synthese*, 195(6), 2541–2557.
- Latash, M. L. (2010). Motor synergies and the equilibrium-point hypothesis. *Motor Control*, 14(3), 294–322.
- McDowell, J. (1979). Virtue and reason. *The Monist*, 62, 331–350.
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 196–204.
- Mele, A. R. (1992). Recent work on intentional action. *American Philosophical Quarterly*, 29(3), 199–217.
- Mohan, V., & Morasso, P. (2011). Passive motion paradigm: An alternative to optimal control. *Frontiers in Neurorobotics*, 5, 1–28.
- Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, 24(5), 776–781.
- Nanay, B. (2013). *Between perception and action*. New York: OUP.
- Nolan, D. (1997). Quantitative parsimony. *The British Journal for the Philosophy of Science*, 48(3), 329–343.
- O'Connor, T. (2011). Agent-causal theories of freedom. In R. Kane (Ed.), *Oxford handbook of free will* (2nd ed.). Oxford: OUP.
- Pacherie, E. (2012). Action. In K. Frankish & W. Ramsey (Eds.), *The Cambridge handbook of cognitive science* (pp. 92–111). Cambridge: CUP.
- Scarantino, A., & Nielsen, M. (2015). Voodoo dolls and angry lions: How emotions explain arational actions. *Philosophical Studies*, 172(11), 2975–2998.
- Schlosser, M. E. (2007). Basic deviance reconsidered. *Analysis*, 67(295), 186–194.
- Searle, J. (1983). *Intentionality: An essay in the Philosophy of Mind*. Cambridge: CUP.
- Seth, A. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, 17(11), 565–573.
- Shea, N. (2013). Perception versus action: The computations may be the same but the direction of fit differs. *Behavioral and Brain Sciences*, 36(3), 228–229.
- Shepherd, J. (2014). The contours of control. *Philosophical Studies*, 170(3), 395–411.
- Shepherd, J. (2017). Halfhearted action and control. *Ergo*, 4(9), 259–277.
- Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience*, 20(3), 458–469.

- Smith, M. (1994). *The moral problem*. Oxford: Blackwell.
- Talsma, D. (2015). Predictive Coding and multisensory integration: An attentional account of the multisensory mind. *Frontiers in Integrative Neuroscience*. <https://doi.org/10.3389/fnint.2015.00019>.
- Tappolet, C. (2010). Emotion, motivation and action: The case of fear. In P. Goldie (Ed.), *Oxford handbook of philosophy of emotion*. Oxford: OUP.
- Todorov, E. (2006). Optimal control theory. In K. Doya, S. Ishii, A. Pouget, & R. P. N. Rao (Eds.), *Bayesian brain: Probabilistic approaches to neural coding*. Cambridge, MA: MIT Press.
- Whiting, H. T. A. (Ed.). (1984). *Human motor actions: Bernstein reassessed*. Oxford: North-Holland.
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, 3, 1212–1217.
- Wu, W. (2011). Confronting many-many problems: Attention and agentive control. *Nous*, 45(1), 50–76.
- Wu, W. (2016). Experts and deviants: The story of agentive control. *Philosophy and Phenomenological Research*, 93(1), 101–126.
- Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: A review of influential factors. *Medical Education*, 44(1), 75–84.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.