

Reviews and perspectives

Why we may not find intentions in the brain

Sebo Uithol^{a,b,*}, Daniel C. Burnston^{c,1}, Pim Haselager^a^a Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, PO Box 9104, 6500 HE Nijmegen, The Netherlands^b University of Parma, Department of Neuroscience—Section of Physiology, Via Volturno 39, 43120 Parma, Italy^c University of California, San Diego, Department of Philosophy, Interdisciplinary Cognitive Sciences Program, 9500 Gilman Drive, La Jolla, San Diego, CA 92093, USA

ARTICLE INFO

Article history:

Received 15 September 2013

Received in revised form

29 November 2013

Accepted 17 January 2014

Available online 24 January 2014

Keywords:

Action

Intention

Prefrontal cortex

Motor control

ABSTRACT

Intentions are commonly conceived of as discrete mental states that are the direct cause of actions. In the last several decades, neuroscientists have taken up the project of finding the neural implementation of intentions, and a number of areas have been posited as implementing these states. We argue, however, that the processes underlying action initiation and control are considerably more dynamic and context sensitive than the concept of intention can allow for. Therefore, adopting the notion of 'intention' in neuroscientific explanations can easily lead to misinterpretation of the data, and can negatively influence investigation into the neural correlates of intentional action. We suggest reinterpreting the mechanisms underlying intentional action, and we will discuss the elements that such a reinterpretation needs to account for.

© 2014 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	129
2. The folk notion of intention and its characteristics	130
2.1. Context-independence	130
2.2. Discreteness	131
2.3. Different levels of intentions	131
3. Action control in the prefrontal cortex	132
4. The complex and dynamic nature of action control	132
4.1. Single cell studies	132
4.2. Imaging studies	133
4.3. Complexity and the folk notion of intention	133
5. Various motivations for assuming discrete intentions	134
5.1. Motivation: We need increasing abstraction to account for increasing complexity	134
5.2. Motivation: There are successful models using discrete states	135
5.3. Motivation: We need intentions for self-initiated action	135
5.4. Interim conclusion: Heuristics, concepts, and mechanistic understanding	136
6. Intentional action without intentions?	136
7. Conclusion	138
Acknowledgements	138
References	138

1. Introduction

Actions are generally thought to be the result of a preceding intention to act. You intend to grasp the cup in front of you, and subsequently (and consequently) you grasp the cup. Upon being asked why you grasped the cup, you will probably reply by stating

* Corresponding author. Tel.: +39 0521 903606/903879; fax: +39 0521 903789.

E-mail address: sebouithol@gmail.com (S. Uithol).¹ The first two authors contributed equally to the paper.

your further intentions concerning the cup, e.g. drinking from it, or putting it in the dish washer. Our daily communicative practices are full of 'intention talk'. By formulating intentions we can describe, explain and predict our own behavior and that of others. This 'folk psychological' use of the concept of intention conceives of intentions as clearly identifiable, relatively simple mental states, free from context-specific details, that are the originating causes of subsequent action planning and motor movement. As such, the notion of intention figures in a variety of contexts, including psychology (Meltzoff, 1995), and philosophical theories of action (Bratman, 1987; Davidson, 1963). More recently, the folk conception of intention has been adopted in neuroscientific studies into willed action. For example, Haggard summarizes the role the notion plays in computational neuroscience as follows: "In computational motor control, for example, actions begin with a relatively simple description of a goal (e.g. 'I will stand up'). The brain must expand this task-level representation into an extremely detailed movement pattern [and] specify the precise kinematics of all participating muscles and joints. Generating this information is computationally demanding. The brain's solution to the problem may lie in the hierarchical organization of the motor system. Details of movement are decided at the lowest level of the motor system possible" (Haggard, 2005).

Attempts to localize the neural correlates of intentions, however, yield diverging results. For example, Lau, Rogers, Haggard and Passingham (2004) asked subjects to attend to their own intentions while performing an action, and found increased activation in the pre-SMA region of the medial prefrontal cortex. Haynes et al. (2007) report finding neural activation specific to subjects' intentions to either subtract or add in the medial prefrontal cortex (more anterior than Lau et al. reported), as well as lateral prefrontal cortex (see Section 5 for a discussion of these findings). As a last example, Carota et al. (2010) report that Broca's area and parietal areas control the intention to speak. If intentions are indeed context-free, amodal and high-level states, as suggested by Haggard's quote above, it is rather puzzling that they are reported to be localized in such wide-ranging areas, which seem to be related to the modality of the subsequent action.

There are also a variety of different reports about the time-course of intentions and their relation to action. Libet (1985) famously showed the existence of an action-related readiness potential 300 ms prior to the reported conscious intention. More recently, Soon, Brass, Heinze and Haynes (2008) were able to predict a decision to make a left or right index finger movement from MRI data up to 10 s prior to the reported time of decision. But if intentions are supposed to be the direct causes of actions, it is not clear why they should exhibit such a broad range of temporal relationships to the actions they are posited to cause.

If each of these investigations is looking for the same thing – namely, discrete intentions – then they produce seemingly conflicting anatomical and temporal localizations for the same type of mental state. We suggest that the disagreement is an effect of the assumption that a discrete state is responsible for the subsequent actions. This assumption is, we believe, based on two ideas. First, the consequences of action control – i.e. actions – also have an apparently² discrete nature. It is therefore assumed that the mechanisms responsible for this output must also make use of discrete states. Second, there is the, perhaps implicit, idea that the way in which we conceptualize intentions in our daily descriptions and explanations of behavior, capture the basic properties of the neurophysiological states implementing them.

² We say 'apparently discrete', because what is taken to be 'the action' in an ongoing stream of behavior is also a matter of interpretation (see also Barker (1963)).

However, a conceptual taxonomy utilized for daily social interaction ('folk psychology') need not offer a valid framework for research in cognitive neuroscience. Moreover, as we will argue throughout, the presence of seemingly discrete outputs from a process does not mean that the process itself contains or involves discrete internal states. In this paper we argue that investigations of intentional action at the neural level should consider intentions as considerably more dynamic and context dependent than our everyday manner of speaking about them suggests. We will discuss evidence that action control, specifically as implemented in the prefrontal cortex, is based on dynamic processes that are continuously updated and deeply sensitive to perceptual and motor context, and we will argue that these characteristics are incompatible with discrete intentions.

Consequently, our view suggests a substantively different approach to a neuroscience of willed action. Instead of using discrete states as part of the explanation of action control, it becomes cognitive neuroscience's project to explain how states that are stable enough to plan and control action can occur in a dynamic and context-sensitive structure. We will briefly sketch the outlines of such an alternative approach. To begin, however, we will briefly discuss the properties of the folk notion of intention, and its philosophical counterparts.

2. The folk notion of intention and its characteristics

The notion of 'intention' plays an important part in our folk psychology, our everyday framework of explaining the behaviors of ourselves and others (Anscombe, 1957; Davies & Stone, 1995; Haselager, 1997; Stich, 1983). There has been intense philosophical debate about the theoretical status of folk psychology in general (e.g. Churchland (1989), Fodor (1987), Greenwood (1991)). Here, we attempt a slightly different approach, by focusing mainly on the supposed evidence for a folk psychological category – intention – and examining whether the empirical data provide genuine support for such a notion. We will therefore not discuss the philosophical literature on the topic in detail. Two characteristics of the folk notion of intention are important for our project: Intentions are generally conceived of as being (1) context-independent, and (2) discrete states. We will discuss these two properties in more detail below.

2.1. Context-independence

Pacherie (2008), Searle (1983), and Bratman (1987) argue that the content of a particular intention is independent of the perceptual, affective, and cognitive context in which it is implemented, and therefore each particular intention needs to be subsequently embedded into a context in order to cause an appropriate action. This is not to say that these philosophers claim that the *forming* of an intention is always independent of a context: a specific context, say seeing fruits in the supermarket, can help bring about an intention to eat an apple (for instance, by causing a desire for an apple). What is meant is that the content of the intention 'I will eat an apple' is the same irrespective of the color or shape of the apples one intends to grasp, or the particular perceptual surroundings at the time or the reason for grasping them.³ This context-independence allows one to form intentions

³ It is also possible that one could encode specific properties of the intended outcome. One might form the intention, for example, to grasp a red apple at the Acme next Thursday. (cf. the 'neurotic planner' in Pacherie and Haggard (2010)). One could, however, form this intention at any time, and it remains unchanged regardless of, e.g., the properties of the apple one eventually decides to grasp, or the room one is in when entertaining the intention. The important point is that

about future actions in a different context, such as picking-up groceries after work (Pacherie, 2008). Consequently, particular intentions retain their characteristics across different instances. If you intend to grasp an apple today, the intention is the same as when you intended to grasp an apple last week, or last year.

2.2. Discreteness

Next, intentions are thought to be *discrete* mental representations. A system implements discrete representations, according to Dietrich and Markman (2003), if it contains “more than one representation, and the representations are bounded and uniquely identifiable.” In philosophical accounts (and, we believe, in neuroscientific contexts), two related forms of discreteness are attributed to intentions, either implicitly or explicitly. First, intentions are thought to have discrete *content*. Content discreteness means that there are clear boundaries between what the state refers to and what it does not. Since content discreteness bounds what the intention refers to, it guarantees that an intention, say the intention to grab an apple, will apply to all and only acts of grabbing apples, thus providing the desired generality across different instances of apple-grabbing, and different contexts (see also Dietrich and Markman (2003)). Also, the content of a discrete intention remains unchanged during the course of an action. In order to keep their content the same, discrete intentions must not be modified by other processes during the ongoing action (although they can perhaps be discarded and replaced by other intentions in light of changing surroundings, for instance if the presence of a more delicious fruit causes the formation of a new, alternate intention, namely to grab the alternative fruit).

Next, intentions are thought of as being *functionally discrete*, which means that they are cognitive units that play a *distinct and identifiable* functional role in the cognitive system (Haselager, 1997; Ramsey, Stich, & Garon, 1999; Stich, 1983). So, the grasping of an apple is caused by the intention to grasp an apple, and the absence of the intention would result in not grasping an apple (except perhaps by accident, or through manipulation). Functional discreteness also means that the state retains its identity while causally interacting with other states—the intention to grasp an apple remains identifiable as that intention throughout the action, even though it may interact causally with the more complex states that cause motor behavior.

Content and functional discreteness are closely related, and the content of an intention is sometimes even attributed based on the role this state plays in the cognitive system—Dietrich and Markman, for instance “define semantic content [...] in terms of the *use* to which that information is put” (ibid., p. 97, our italics). For example, a state can be said to represent the intention to grasp an apple (content) when it makes us grasp an apple in the appropriate situation (function). More generally, Pacherie (2011) claims that “intentions can have a causal impact on action by *virtue of their contents*” (p. 67, our italics) suggesting a tight coupling between function and content. Finally, it is essential to explanations and predictions of behavior to indicate *which* intentions play the causally crucial roles in the production of behavior. Hence content and causal role of mental states need to be tightly coupled.

Functional and content discreteness do not entail context-insensitivity, but are related to it. For example, Pacherie (2008, p. 183) specifically posits that it is the *abstractness* and context-insensitivity of intentions that allows them to play their functional role, and discrete states are well-positioned to exemplify this type

(footnote continued)

specific perceptual and motor elements of the actual action context are not represented in the intention.

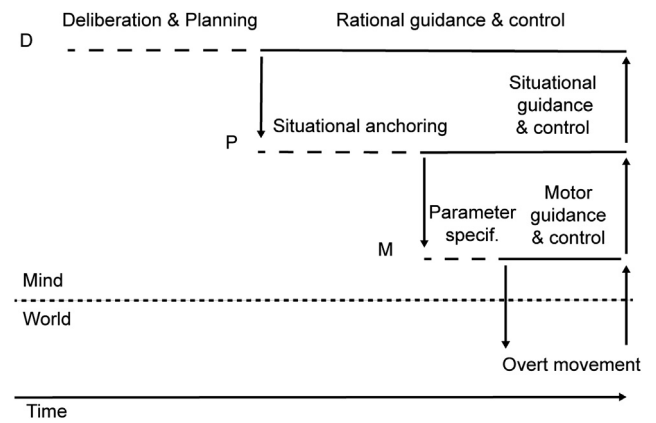


Fig. 1. An overview of Pacherie's “causal cascade” theory of intentional action, taken from Pacherie (2008).

of property. Since they are not modified by interactions with other states, they can be dissociable from contextual details (i.e., distinct from other states that represent those details) and employed in multiple different contexts.

2.3. Different levels of intentions

Representations of actions are usually thought of in a hierarchical fashion. On top of the hierarchy one can find future-directed and context-free intentions involved in long-term planning of actions that consist of many sub-actions, whereas ‘mid-level’ states involve aspects of the current action context, and at the bottom states represent the specific movements or actions.⁴ Generally, the different representations are thought to stand in causal relationships to each other, with the higher more future-oriented intentions being the cause of the contextual influences and motor states, forming a hierarchy of action control (see also Grafton and Hamilton (2007), Haggard (2008) for the virtually identical ‘action hierarchy’; Hamilton and Grafton (2007)).

Pacherie (2000, 2008) has done admirable work to make this general framework empirically tractable (see Haggard, 2008; Hamilton & Grafton, 2007; Moore, Wegner, & Haggard, 2009; Ondobaka, Lange, Newman-Norlund, Wiemers, & Bekkering, 2012; Pezzulo & Castelfranchi, 2009; Wenke, Fleming, & Haggard, 2010 for examples). In her framework (Fig. 1), the most context-free states are at the top of the hierarchy, and are important for deliberation and rational control, while all perceptual context and motor associations take place lower down in the hierarchy. The higher-level intentions capture the intuitions present in the folk interpretation of intentions (that they are context-insensitive, discrete and the originating cause of actions), while the lower-level intentions are supposed to provide the bridge between these abstract and deliberative processes on the one hand and physical movement on the other.

In what follows, we will argue that, despite the intuitive appeal of the picture, intentions as discrete states that cause our actions should not be assumed as an explanatory construct when investigating the neural mechanisms for action control. We will discuss evidence that the putative function of intentions – that of

⁴ Many philosophers have accounted for this difference by distinguishing different types of intentions. Searle (1983) contrasts *prior intentions* and *intentions in action*; Bratman (1987) uses *future-directed* and *present-directed* intentions, Pacherie (2008) contrasts *distal*, *proximal* and *motor* intentions, and Pacherie and Haggard (2010) use *prospective* and *immediate intentions*. To our understanding, they roughly denote the same distinction. Since our main target corresponds to the context-free type of representation, we will reserve ‘intention’ for context-free states.

underlying temporally extended, deliberative action control – is implemented by the anterior areas of the lateral prefrontal cortex (IPFC), and that the neural processes that shape action control in the PFC do not exhibit the typical characteristics of intentions that we have discussed.

3. Action control in the prefrontal cortex

Based on a variety of evidence from imaging studies, neurophysiology, and neuropsychology, the prefrontal cortex (PFC) is generally recognized to be a primary locus of action generation and control in the brain (see Müller and Cohen (2001) for an early review of this evidence). Various models of prefrontal action control have been proposed (see Badre (2008) and Ramnani and Owen (2004) for reviews), and while the details of different models of control differ, the notion of an anterior–posterior axis within the lateral PFC (IPFC) for implementing different kinds of control is now widely recognized (Badre, 2008; Fuster, 2004).

Action generation processes in the IPFC, especially in anterior regions, seem to exhibit the same temporally extended and deliberative aspects as intentions. For example, Koechlin and Summerfield (2007), Koechlin, Ody, and Kouneiher (2003) have suggested and tested a model in which different types of action coordination are subserved by different areas of the IPFC. Based on differential activation found in imaging studies, they posit that posterior areas of the PFC are in charge of selecting between different “sensorimotor associations,” where these consist of a learned connection between a stimulus property and a specific motor act. When different sets of context-dependent associations must be applied depending on the nature of the stimulus sets presented, dorsolateral PFC (dlPFC) is activated, performing a process that they call *contextual control*. When these contextual processes are sensitive to certain past events (e.g., which stimuli have been perceived previously), this is referred to as *episodic control*. Lastly, when a task must be paused, and a new set of associations implemented, due to the presentation of an interrupting stimulus, followed by a return to the first task, the lateral frontopolar cortex (the most rostral part of the PFC) is activated, a process they call *branching control*.

Like intentions, the processes in episodic and branching control are temporally extended, in the sense that they must keep track of information longer than other types of control—they involve maintaining representations of a variety of task rules in order to act appropriately in response to the presented stimulus. They also seem to be genuinely deliberative, since they must keep track of a variety of different considerations, and discern the appropriate relationships between them, in order to complete tasks effectively.

These studies suggest that the functions attributed to the regions along the anterior–posterior axis seem to be similar to the functions of intentions and the several other layers of the hierarchy. The posterior prefrontal cortex is thought to be involved in concrete action responses (Badre, 2008); More anterior regions are progressively involved in more temporally extended and deliberative action planning (Badre, 2008; Christoff & Gabrieli, 2000; Koechlin et al., 2003), just as intentions are supposed to be.⁵ So, in behavior requiring branching control, anterior areas would store the required future associations, as well as the contexts in which they need to be employed, and initiate processes to apply the new rules in the appropriate setting. Koechlin et al. (2003) suggest that a broad “goal representation” is maintained in aPFC during the performance of sub-tasks, and activates the appropriate

actions at the appropriate times. Burgess, Veitch, de Lacy Costello and Shallice (2000) interpret the deficits in patients with lesions in aPFC as an inability to form and carry out intentions, and Ramnani and Owen (2004) describe the relevant sort of intention along the lines of such plans as to ‘Meet John at 5’: an obvious parallel to future-directed aspect of intentions.

In all, if intentions are to be found in the brain, the most likely location would be along the anterior–posterior axis of the IPFC. However, despite the similarities between functions attributed to the intentional hierarchy and to the processes along the anterior–posterior axis, we will argue in the next section that the context-independent, discrete, and causal rendering of intentions is incompatible with the type of information processing that occurs in the PFC.

4. The complex and dynamic nature of action control

Intentions, in the folk and philosophical senses, are essentially simple, discrete, and independent of the context in which the actions they cause are embedded. In this section we will show that there is convincing empirical evidence that neural activity in the anterior regions of the IPFC is *informationally* as well as *dynamically complex*. These types of complexity suggest that processing in the IPFC does not exhibit the type of structure that is attributed to intentional action generation when looked at from a folk psychological perspective.

Let us start with informational complexity. A system's operations are informationally complex, in our view, if during its normal operation the system has access to and makes use of a variety of different sources of information. The anatomy of the IPFC suggests that it has access to a variety of different sorts of information. The anterior regions of the IPFC receive input from the medial temporal lobe and the thalamus, as well as multiple sensory areas, such as ventral visual areas, somatosensory areas, auditory cortex, and the rostral superior temporal sulcus, which is itself known to be a multimodal area (Miller & Cohen, 2001; Ongür & Price, 2000). While anatomical connections themselves do not impel conclusions about the nature of processing, this connectivity at least suggests that the resources for informationally complex processing are present. Direct evidence for informational complexity comes from physiological studies at multiple spatial scales. In the next subsections, we discuss a few specific studies that make these points in detail.

4.1. Single cell studies

Some early single cell studies by Fuster, Bauer and Jervey (1982) showed evidence that, in monkeys, anterior dorsolateral and orbital regions of the PFC had specific responses to perceptual features of the action context, and to the relations between them. Fuster et al., conclude that these data “provide [...] evidence for the hypothesis that the prefrontal cortex is essential for the temporal integration of sensory data and motor acts in sequential behavioral structures” (p.690). Following on these results, Fuster, Bodner and Kroger (2000) trained monkeys to form an action-related association between tones and colors. Specifically, the monkeys had to press a certain color on a touch screen following the presentation of a certain frequency sound. Single-cell recordings were conducted in over 300 neurons spanning areas in the dorsolateral and anterior prefrontal cortex. In accordance with the previous findings, they discovered that while some cells encode the presence of only the tone or only the color, a large quantity of the cells increased activation in the interim between the presentation of the tone and the colors, suggesting that these encode the presence of a relation between tone and color.

⁵ Pacherie and Haggard (2010, p. 74) cite the dlPFC as being involved in “selecting actions.” This could be interpreted as the translation of future-directed intentions to specific motor intentions.

Fuster et al. (1982) contend that the function of these PFC cells is to facilitate the *integration* of perceptual information and, moreover, to do so *over an extended period of time*. The cross-temporal aspects of these neurons' activity suggests an action control process that seems to be in line with the cross-temporal role that intentions are supposed to play; however, the process seems highly sensitive to specific perceptual information. It is the presence of *the associated stimuli* in the perceptual environment that mediates processing in the LPFC, and this processing is directly relevant to performing the proper action (via the learned perceptual association).

One could maintain that the cells that associate a tone and a color are encoding a discrete intention, such as "I will press red." This would be to invoke the folk notion that perceptual information is turned into a context free state, for instance in the neurons which respond solely to tones (e.g., "the tone is present"), which in turn is formed into a discrete intention to press the appropriate color. However, the commitment to discrete states suggests a stable firing pattern for the duration of the action. This commitment is directly undermined by the *dynamical complexity* of the processes that Fuster et al. discuss. We define dynamical complexity as the continuous modification of the state of a system in accordance with processing and task demands. In accordance with this definition, the neurons that Fuster et al. (2000) studied which showed feature-specific associative behavior, showed distinct behaviors at different states of the action sequence. Some neurons maintained correlated firing only through certain epochs of stimulus presentation – for instance, during auditory presentation and in the delay between stimuli – while others fired consistently in different stages. Moreover, even the neurons that seem to represent the association between the stimulus features, as opposed to neurons that fired for just one of the features – i.e., the ones that were discussed above as candidates for encoding the discrete intentions – did not demonstrate a constant response, but instead their activity varied in their temporal relations to the onset stimulus. So, in contradiction to the conditions on positing discrete representations, the activation of the neurons studied by Fuster et al. do in fact vary continuously during the period of action control, and moreover in a way that is sensitive to the fine-grained temporal progression of the action sequence. In order to maintain discrete, context-insensitive content for these states, the discrete theorist seems to be forced into the position of asserting that these temporal characteristics of activity have no functional role.

4.2. Imaging studies

The properties discussed above are in line with findings of imaging studies that suggest the presence of informational complexity at the systems-level for all regions of the human LPFC. For instance, more posterior areas of the LPFC have been implicated in contexts where associations between visual stimuli and particular actions (Passingham, Toni, & Rushworth, 2000) need to be recalled (see also Koehlin et al. (2003), Kounieher, Charron, and Koehlin (2009), Rushworth (2008)). This suggests that activity in these areas is modified depending on whether there is perceptual information needed to complete a task. Importantly, these effects are not restricted to posterior regions of the LPFC, as would be expected if representations became less context-sensitive as one moved anteriorly along the anterior–posterior axis. Prabhakaran, Narayanan, Zhao and Gabrieli (2000) compared two conditions involving tasks cued by verbal and spatial information, one in which the two types of information were presented separately, and one in which the verbal and spatial cues coincided. In the "non-integrated" (first) condition, activation was higher in more posterior areas of the LPFC, whereas in the "integrated" (second)

condition, activation was considerably higher in more anterior regions, particularly in the right hemisphere.

The difference between posterior and anterior regions, then, is not in *whether* information about perceptual context is processed in anterior areas, but *what kind* of context information is processed – i.e., information about specific environmental features more posteriorly, and information about relations (e.g., spatiotemporal coincidence relations) between environmental features more anteriorly (see also Christoff, Ream, Geddes, and Gabrieli (2003)). Similarly Ranganath, Johnson and D'Esposito (2000) found that the left anterior LPFC shows increased activation during tasks where more specific perceptual information had to be recalled, suggesting not only that informational complexity increases as one moves more anteriorly along the LPFC axis, but that the overall *quantity* of perceptual information that is needed for a task modifies aPFC activity. These results strongly support informational complexity in the LPFC. While dynamical complexity is difficult to show with techniques such as fMRI, the Ranganath et al. (2000) study suggests that the amount of perceptual detail encoded by these areas can shift with task demands. Importantly, the evidence ranges across spatial scales, from individual groups of neurons to the activity of whole regions.

4.3. Complexity and the folk notion of intention

The types of complexity discussed above are difficult to reconcile with the three properties of the folk notion of intention discussed in Section 2; context independence, discreteness and action causation. First, *context-independence*: The informational complexity of the PFC shows that even the most anterior parts of the LPFC are not context-insensitive in the way demanded by the core folk notion of intentions. It is problematic to frame the representation of higher-order relations (e.g., spatio-temporal ones) between perceptual information in terms of discrete states. Indeed, if the role that Fuster et al. (1982), (2000), Fuster (2001) attribute to LPFC neurons of integrating perceptual information over a period of time is correct, then this context-sensitivity is *vital* for the functioning of these areas. Moreover, dynamical complexity suggests a continual dependency on perceptual context. The view suggested by these results is that branching and episodic control (Koehlin & Summerfield, 2007) implement their functions by tracking information relevant to a particular task or task set in the actor's perceived environment over time. Branching control, for instance, attempts to mediate shifting task demands (Collins & Koehlin, 2012; Rogers & Monsell, 1995; Sakai, 2008) with changes in context, and does so by keeping track of relations between a variety of different informational sources. On the view suggested by informational and dynamic complexity, this sort of procedure can produce temporally extended behaviors without relying on previously represented, discrete intentions, through dynamic interaction with a perceptual context. The prospective memory results also show, importantly, that perceptual context need not be *occurrent*—it can be recalled or expected context that allows for the future-directed aspects of actual action.

Second, the two kinds of complexity also undermine the notion of intentions as *discrete* states. Dynamical complexity alone would be problematic for an attempt to isolate a discrete state from ongoing processes, since (if Fuster et al.'s results generalize) it shows that activation in anterior regions of the PFC is continuously modified from the earliest stages of the task (certain aspects, however, do remain specific across changes in task context; we discuss these in Section 5). This speaks against the idea of a discrete state that begins the action coordination process, and retains its functional identity and content throughout and action episode. The additional fact that these processes are embedded in context up to the highest levels means that it is very unlikely that

one will be able to isolate a state whose processing or content is dissociable from the context in which it occurs. Of course, some elements of the context may remain similar across instances of apple-grasping (e.g., apples tend to have roughly the same shape and size across instances), and therefore some elements of the apple-grasping process may be similar across most instances (e.g. grip aperture). However, this does not mean that there is a state somewhere in this process that is discrete and dissociated (contentfully and functionally) from context. As Fuster et al. (2000) suggest, these shared aspects can simply be due to the influence of contextual elements that are similar across those instances, thus arguing that any apparent discreteness here is a by-product of context-sensitive processes being employed in similar contexts.

The combined results are incompatible with the general view of the action-hierarchy in which context-free states sit at the top and cause particular activations of context-sensitive representations. The informational and dynamical complexity of IPFC control processes undermine the idea of increasing abstractness or simplicity as one moves up to more deliberative and temporally extended aspects of actions. Instead, the data discussed above suggests that there is a different *kind* of complexity present at higher levels. At low-level motor control there is complexity with respect to detailed information regarding motor states – e.g., proprioceptive and sensorimotor feedback from the muscles involved – and detailed plans for how to perform suites of motor effector movement. At higher, deliberative, levels, there is a complexity of *types* of information (e.g., information from multiple senses, as well as current motor-state, occurring across extended timescales) and the relations between them. So unlike low-level control, in which there is complex and detailed information *within* a type of information (e.g., effector-specific, or within a single perceptual-motor mapping), higher levels of control are complex in that they can process relations between a variety of stimulus types, rules, outcomes, and temporal contexts. This is far removed from the folk notion of a simple and discrete state.

We have discussed these results specifically and in detail, because they straightforwardly exhibit the properties of informational and dynamical complexity. We take them to be entirely emblematic; in general, the evidence from single cell and imaging studies shows that the prefrontal cortex is sensitive to perceptual information at all levels, and is involved in flexibly mediating this information in response to context and task demands (Genovesio, 2006; Genovesio, Brasted, Mitz, & Wise, 2005; Mushiake, Saito, Sakamoto, Itoyama, & Tanji, 2006). In the next section, we discuss a variety of standard empirical and intuitive motivations for positing discrete intentions, and argue that none of them provide genuine reason for doing so.

5. Various motivations for assuming discrete intentions

So far we have suggested that action control in the PFC – including its most anterior regions – exhibits informational and dynamical complexity, and that these properties are incompatible with intentions playing a key causal role in action generation. Yet, intentions are often assumed to be necessary to explain a wide range of empirical results. In this section, we discuss a number of motivations that could lead to the assumption of such discrete states underlying action control, and we discuss some of the empirical data that is interpreted using such assumptions. We show why in our view this work need not be considered to be evidence for discrete states, and how it can also (and perhaps better) be interpreted without such assumptions. Throughout, we will also briefly discuss alternative evidence that speaks against interpretation in terms of discrete intentions.

5.1. Motivation: We need increasing abstraction to account for increasing complexity

One compelling intuition is that increasing complexity of control goes together with increasing abstraction, and decreasing context-sensitivity. Many results in the action control literature assume that the structure of tasks is itself hierarchical, with goals at the top of the hierarchy, and sub-goals nested within it. It is also often assumed that the hierarchical organization of the control system must match that of the goal hierarchy (Uithol, van Rooij, & Bekkering, 2012). Such a hierarchical organization brings with it the intuitive assumption that the higher levels of the control hierarchy operate over more abstract representations, perhaps ones that are entirely free of context. To a certain extent such a perspective is encouraged by the model of Koechlin et al. (2003), Koechlin and Summerfield (2007) that we have leaned on in our characterization of the PFC. The addition of different levels of control involves nesting a set of conditional dependencies between sensorimotor associations—branching control, for instance, involves interrupting sensorimotor associations in response to new stimuli, while episodic control makes these cue-action associations dependent on a background condition which varies per stimulus block. Thus, the levels of control are themselves hierarchical. This, however, need not involve more abstract and amodal types of representations on top of the hierarchy. Higher-level processes can very well consist of more complex types of associations between more elements of perceptual and motor context.

To give one example, Badre and D'Esposito (2007) varied task structure in ways roughly analogous to the control structures posted by Koechlin et al., and also varied complexity within each control type. They also found that activity increased anteriorly with more complex task structure, and moreover that the particular regions activated in each type of control varied with increased complexity within the type. They interpret this as evidence for abstract representation. However, it is questionable that such an interpretation follows from the results. In the increased complexity conditions, it was exactly an increase in the number of, and nesting relations between, sensorimotor associations that was varied. That is, as the complexity of contextual cues that need to be remembered goes up, so does PFC activation (at each level of control). One would expect that if an abstract, context-free state was being implemented, then activation would not change so drastically with the amount of context tracked.

A related type of interpretation posits that the PFC, specifically more anterior regions, represents task 'rules'. Rules are similar to abstraction in that they are supposed to be distinct from the material to which the rules apply—e.g., a representation “when cue *X* occurs, I must perform task *Y*,” needs to abstract away from both *X* and *Y*, and the similarity of this kind of representation to linguistically phrased intentions is obvious. The primary evidence taken to show such rule implementation comes from situations in which the perceptual stimulus stays the same, but distinct activations can be found for different tasks involving that information. Both imaging (see Sakai (2008), for a review) and single cell results are often interpreted this way, but we will only discuss single cell results for brevity. Genovesio et al. (2005), for example, found neurons that respond preferentially to different tasks involving the same contextual information. Moreover, the activity of these neurons can be initiated before the stimulus is shown, can show activation even if the particular stimulus-response mappings for a task are variable, and can be used to predict which task is performed.

However, the fact that different neurons can be active in different tasks involving the same perceptual context is not proof that a discrete intention is being encoded. As Sakai (2008, p. 226) notes, it is not surprising that different *organizations* of perceptual

information should be mediated by different neurons, since different tasks will demand that this information is grouped in assorted, specific ways. To see this clearly, take a simplified example. Suppose that we have two populations of associative neurons like the ones discussed in regard to Fuster et al.'s (2000) study above. Now suppose that these associations must in turn be associated in different ways, as in more highly nested forms of action control. Suppose that population 1's association must come before population 2's in some contexts, and afterwards in others. Clearly, something must be distinct to implement the two different orderings, but this need not be an abstract rule. Instead, distinct coordinations can be implemented by any mechanism that imposes an ordering between the two populations—for instance, different individual or populations of inhibitory interneurons.⁶

In short, implementing rule-like behavior does not entail the formation of an abstract, discrete state, which is then applied to contextual ones that are functionally distinct from it. Rather, this behavior can emerge from detailed coordination of already existing contextual representations, by further developed organization in particular behavioral contexts. Moreover, the possibility that this organization could even generalize to similar stimuli, or be primed in task anticipation, is compatible with the idea that *what* is being primed is a way of grouping perceptual elements. In other words, the function implemented by the rules involves mapping certain elements of stimulus context onto motor action, leaving no need to posit independent, discrete, abstract states for these effects.

Moreover, other studies using roughly this paradigm can be construed as providing evidence *against* discrete intentions. Simons, Schölvinc, Gilbert, Frith, and Burgess (2006), for instance, distinguish between what they call 'cue identification' and 'intention retrieval'. In cue identification trials, subjects had to only press a button different from the buttons being pressed in ongoing tasks, while on intention retrieval trials subjects had to perform an additional and remembered action (e.g. count the syllables in words and press different keys if the total is less or more than 4). As in the 'rule' results discussed above, in this study the type of operation was different in the two conditions, while the perceptual input was the same. They report that although these two processes can be contrasted on the behavioral level, they share a common neural basis in anterior prefrontal cortex. In this study, 'intention' is operationalized as increase in demand on internal storage (see the discussion of heuristics, below), and the study does not motivate the positing of abstract, discrete, context-free states. When the different tasks are seen as continuous, but relying on a different, and perhaps more complex ordering and grouping of the stimuli, as operationalized in this study, the shared neural substrate is to be expected.

In line with this, Gilbert et al. manipulated how internally generated a task is by making the instructions to perform a previously learned task implicit and abstract rather than concrete and descriptive. The authors showed a slight preference in lateral aPFC (versus medial) for a more internal condition. However, like Simons et al., they reasonably interpret this as a "continuum" between the mechanisms for the two conditions, rather than a categorical difference.

5.2. Motivation: There are successful models using discrete states

Intentions have played a role in modeling in psychology as well as in neuroscientific research, and it might be thought that a successful behavioral model employing discrete states motivates

the search for discrete neural realizers of those states. For instance, Collins and Koechlin (2012) build discrete higher-order task options, or "behavioral strategies," that determine the particular actions (in this case, button presses in response to individual stimuli), into their model of action generation. They show how a model with distinct, explicit strategy representations can fit a variety of behavioral data.

We suggest that the success of such a model at the psychological or computational "level" cannot be interpreted as evidence for discrete states implemented in the neural control system. Any successful psychological model that posits a discrete internal state shows that, in the behavioral contexts for which it is successful, the neural control system *as a whole* must end up at an *end state* that matches the output suggested by the psychological model (e.g., a specific motor action consistent with a particular strategy). Yet this does not entail either functional discreteness or a single abstract state representation in the process leading up to that end result.

As an illustration, consider a range of models of working memory dealing with sensorimotor associations (Christoff, Keramatian, Gordon, Smith, & Mädlar, 2009; Loh & Deco, 2005; Loh, Pasupathy, Miller, & Deco, 2008), in which dynamically interacting neural networks that are completely interconnected end up settling at discrete outputs. On such models, both incoming sensory information and background behavioral context are taken to *bias*, in a probabilistic sense, the eventual choice of sensorimotor association that the network will settle on. Such networks can switch flexibly between discrete outputs depending on different elements of context, and can learn new associations based on changes in input and reward, despite not being controlled by some higher-level, more abstract state, and despite not having functionally discrete states.

So far, these models have focused mostly on low-level sensorimotor associations. However, there is nothing that bars extending such concepts to more complex elements of contextual information and action control—even up to and including the notion of behavioral strategy or task set. These models are clearly distinct from a hierarchy on which an abstract intention sits at the top and determines the processing of lower levels, and posit, via the biasing relationship, a kind of context-dependent action organization that seems more compatible with the context-sensitivity of PFC processing we have discussed. This is not to say that it will not be a significant challenge to understand how these (in our view) more neurally plausible models can match the outputs of high level psychological models. But there is no conceptual barrier to the challenge, and given the current state of the neural evidence, we suggest that it may well be a fruitful research direction.

5.3. Motivation: We need intentions for self-initiated action

Above we showed that even cases of seemingly abstract or rule-dependent actions can be explained by representations that serve to coordinate multiple distinct elements of already presented or expected context. But what if the action is generated not in response to any external cue, but instead due to a 'free' decision. Does this suggest that a context-free intention must exist to cause the subject to act in a certain way?

Haynes et al. (2007) used decoding methods on fMRI measurements to find which areas of the PFC predicted 'freely chosen' actions, in this case whether to add or subtract two numbers. While they found a variety of differences in activity, the greatest predictor was anterior mPFC, from which 71% of eventual choices could be predicted. In this case, there seems to be no current perceptual context on which to operate. Soon, Hanxi, Bode and Haynes (2013) show that these decisions can already be detected above change well before the subject's reported onset of the

⁶ We do not intend to commit to this particular mechanism. In general, so long as the aPFC has a way of ordering distinct sensorimotor associations, then our strategy here will be feasible. For evidence that populations of PFC neurons order subgoals or stimuli in time, see Genovesio (2006) and Mushiaki et al. (2006).

conscious intention. Such results have been reported extensively using a variety of different methods, and have often been interpreted as evidence that conscious intentions are not the cause of our actions (Haggard, 2011; see for example Soon et al. (2008), Wegner (2003)). This line of thinking is in our view a step in the right direction, but we propose to take one further step: in addition to questioning *conscious* discrete states, many of the results we have discussed motivate questioning the causal influence of discrete states themselves.

We have argued above that for a state to be discrete it has to be stable across contexts, and it has to be stable during interaction with other states. We suggest that many of the cases presented as evidence for the discrete state interpretation rest on the absence of context in the experimental setup—there are no contextual features that affect the decision, and the numbers to be added or subtracted are not presented until after the decision has been made. Given the two-option forced choice nature of the decision, however, it is not possible to determine that the separable states represent the distinct decisions discretely, or play the direct causal role attributed to intentions. Any two distinctly presented options will have *some* noticeable difference in the brain's response. But this distinct difference might be due to the experimental setup, rather than showing that there are inherently context free representations. Showing that a state is an 'intention' in the folk or hierarchical sense involves showing that it not only is distinct in discretely presented scenarios, but that this distinction plays a discrete causal role in the system (i.e., that the distinctness underlies discrete representations that play the causal roles attributed). Doing so might not be easy, as context would have to be varied to show this discrete influence (as per the above). But when context is varied, PFC processes, as we have discussed, tend to reflect contextual influence.

An interesting alternative suggesting how self-initiated actions can come about without discrete intentions is provided by Schurger, Sitt and Dehaene (2012), who modeled the readiness potential as found in the Libet paradigm, using a leaky integrator, stochastic noise and a threshold. The rationale behind this model is that both the 'what' and the 'whether' of an action in the Libet paradigm are fixed. Combined with the instruction to 'act whenever you feel like it' these experiment specifics make the exact moment of action execution highly influenced by random noise in the motor system. In a sense, the 'decision' to act in this setup is the crossing of the threshold, which causes the action to be executed (i.e. the readiness potential is not caused by a decision to act, it reflects the decision to act).

What the model shows is that a *detectable* predictor of an outcome can arise, and look discrete, from within an entirely continuous system. Such a result could occur in an MRI study if, for instance, a dynamically interacting network eventually pulls a particular brain circuit (in this case, the mPFC), into a certain interpretation (activation pattern) rather than another (compare the dynamical decision models we discussed in the previous subsection). These results show that there is no compelling motivation for interpreting this interaction in terms of discrete causal influence of one part of the system on the others, as the discrete intention model.

Schurger and colleagues explicitly only modeled the 'when' of an action decision, but similarly continuous process could very well underlie the 'what' and the 'whether' of an action.⁷

⁷ An interesting result perhaps bridging these perspectives comes from Momennejad and Haynes (2012), who had subjects perform a complex, perceptually cued task while internally keeping track of a time delay, after which they had to switch to a different task using the same information. The authors used decoding methods to assess PFC locations at which "what" new task had to be performed and "when" to perform it could be discerned from activity. They discovered that during the delay, medial aPFC reflected the "what" of the future task, while lateral aPFC reflected the "when," but right before initiating the task change, this reversed, with

The bottom line – again – is that what can be discretely contrasted on the behavioral level need not rely on an equally discrete control mechanism. And as we discussed in relation to behavioral models, such decoding results can serve as a data point and a useful target for more dynamical perspectives, without forcing an interpretation along discrete, hierarchical lines.

5.4. *Interim conclusion: Heuristics, concepts, and mechanistic understanding*

In our view, discreteness is more apparent than actual. The way experiments are set up can make things look discrete, especially by generalizing in specific ways over certain outcomes. However, it may well turn out to be that what looks like a discrete state will appear to be less discrete upon a more detailed inspection. We are not denying that there are descriptions possible that construe action control processes as discrete; nor are we denying that these descriptions could have scientific or heuristic value. In scientific practice, it is often convenient or even necessary to employ localization heuristics in initially decomposing a complex system in order to gain tractability in understanding that system. In neuroscience, claims about representational capacities are often used in this respect (see Bechtel (2003, 2008) for an analysis of this explanatory strategy in a variety of different neuroscience contexts).

Such heuristic use of a concept can aid investigation by directing attention to a certain brain region, and by providing general decompositions of systems from which to work. However, as discussed in the introduction to this section, heuristics when taken too literally can overly constrain thinking about the mechanisms one is studying. We have pointed out several cases in which all of the hypotheses are framed in terms of intentions. If these are being used literally, then they face the empirical and conceptual problems for intentions that we have discussed. If they are being used as heuristics, then attention must be paid to how to move past them into a deeper understanding of action control. How we think about the status of the concept really matters. For instance, Haynes et al. (2007, p. 324) consider their result important for thinking about other executive processes, such as "attending to and thinking about intentions," as well as a variety of other control tasks. Should we pursue these questions as though a discrete control signal is propagated from mPFC (or, speaking more generally, anywhere where intentions or rules are posited to be represented), or in terms of the informationally and dynamically complex interactions we have discussed? We have suggested that the empirical evidence speaks in favor of the latter, and at the very least researchers should be clear on the roles for which they employ such multifarious concepts as 'intention'. In the last section, we give a brief outline of how one might approach action control without giving intentions a primary role.

6. Intentional action without intentions?

We have argued that we should not assume that discrete states underlie action control, despite apparent discrete output. The neural data that we have discussed suggests that the idea of a discrete intention causing and controlling actions from the top of a representational hierarchy is a mischaracterization of the complex and dynamic nature of action control. The causal relationship between a discrete intention and action, and their assumed

(footnote continued)

medial areas predicting "when" and lateral predicting "what". In combination with the Simons et al. (2006) results we discussed above, we find this to be indicative of dynamical complexity, while still involving information localized at particular times and areas. Much more work, of course, needs to be done to understand the details of these processes.

context-independence are all difficult to reconcile with the informational and dynamical complexity of action planning and neural control processes found in the IPFC.

In our view, these neural data suggest the need for a reinterpretation of intentional action at the neural level. We cannot present a complete theory in this paper. Instead, we will discuss the elements that such an alternative interpretation of action planning needs to account for, or could incorporate, and how doing so can shape further investigation into action control at the neural level.

The neural data we have discussed suggest that action control is accomplished by various interdependent control processes, in continual interaction with contextual features. A variety of perceptual relations are represented in control areas, depending on the specific needs of the action context, where these needs are dynamically changing. The resulting control of behavior thus emerges from the interaction of the task, the various processes with each other and with the action environment. With this in mind, it is easy to diagnose why both spatial and temporal localization projects have come up with such diverse results—the situation stems from attempting to apply a single category to a set of processes that are dynamically driven by task and context.

The difficult problems to solve for a future account of intentional action involve both how to individuate control processes, and how to explain their interactions. Which tasks can rely on faster and more direct processes, and which tasks require additional control processes? In what way do these additional processes influence the more direct ones? Several models of action control in the PFC, including some that we have discussed (Christoff & Gabrieli, 2000; Courtney, 2004; Koechlin et al., 2003; Petrides, 2005), have already begun to make progress on these questions. Here we give a few ideas on how these models and findings could be integrated into an alternative account of intentional action.

Although the details and approaches of various models differ, they seem to converge on the idea that the more rostral an area, the more informationally complex and temporally extended the control processes it exhibits. Since different elements of action control instantiate different relations between perceptual information, a complete account of intentional action will have to explain how both stable action features short term motor control emerge from different relations between distinct sorts of perceptual information (see also Cisek (2007), Cisek and Kalaska (2010), Kiebel Daunizeau and Friston (2008), Uithol, van Rooij, Bekkering and Haselager (2012) and Yamashita and Tani (2008)).

Consider the problem that a control system has to solve in the context of complex action planning, say finding the right type of spaghetti at the supermarket. A variety of goals and subgoals – with their concomitant perceptual cues and motor tasks – need to be coordinated in such a way that the distal temporal end-state is reachable via a series of successive previous states, each of which must be ordered in relation to each other in space and time. Moreover, the system must be sensitive to changes in the external environment. A problem at a given stage necessitates the re-coordination of an indeterminate number of intermediate stages, which must be adjusted flexibly, accurately, and perhaps quickly, in order to reach the end-state. In addition, unforeseen intervening stimuli must be judged to be relevant or irrelevant, and the correct ordering of spatial and temporal sub-tasks must be maintained despite distracters.

The hierarchical models by Koechlin et al. (2003), Koechlin and Summerfield (2007) and others, suggest that more posterior areas of the control network are more limited in the types of perceptual associations they can implement, and in their degree of flexibility. Thus, posterior areas might be in charge of looking for particular elements of the action context that are relevant to achieving the

goal (e.g., picking the right package from the shelf), higher areas might be involved coordinating the large number of these spatially-indexed subgoals, and ordering them in space and time, such that one can efficiently walk down the pasta aisle, avoiding distractions, while still procuring all of the ingredients needed for dinner.

Even if some detail is lost in this process, this *does not entail* that contextual information is not relevant or processed. Once again, on the view suggested by informational and dynamical complexity, higher levels of action control involve broader (both spatially and temporally) and more flexible associations between elements of perceptual and motor context, not an abstraction away from all elements of that context. The amount of perceptual detail at which these representations operate might even vary given task constraints, comporting nicely with the data of Prabhakaran et al. (2000) discussed above.

One way to create the flexible associations necessary for action control might be found in Hommel's (2004) 'event files,' which are multi-layered networks of bindings that temporarily link perceptual events, the current task context, and a possible action. The flexible and heterogeneous composition of these files provides the flexibility and context-dependency that is needed for action control, but that is lacking in the notion of intention. An existence proof for the possibility of such emergent or implicit control structures can be found in the modeling work by Botvinick and Plaut (2004, 2006). Despite the absence of explicit processing hierarchies, these models were able to successfully simulate human action planning.⁸

As briefly mentioned above, another way to craft the action control structure is using the notion of 'task set' (Sakai, 2008). A task set, in Sakai, 2008's p. 232 phrasing, is "a configuration of perceptual, attentional, mnemonic, and motor processes"—that is, the neural network involved in implementing a specific task involves a number of brain areas operating in parallel to coordinate the perceptual and motor processes relevant for the task. If the processes included in a task are viewed as in continuous and dynamic interaction with perceptual and motor processes, and not attributed the role of primary initiator of an action, the notion of 'task set' becomes far more compatible with the informational and dynamical complexity of the neural processes we have discussed above. We should not, however, fall into the same trap as with intentions, by assuming that these 'task sets' must embody contentfully and functionally discrete states, or that the mechanisms underlying them can always be contrasted discretely. Even if they result in stable outcomes, task sets themselves may be comprised entirely of context-dependent and dynamic processes.

As stressed, this is not a stand-alone theory of action control, but a perspective that might be taken towards research into the neural processes underlying intentional action. The proposed alternative perspective shifts attention away from localization, as one would expect a control system of the type outlined above to be in continual and various interactions with both perceptual and motivational information. Here, however, is where research is most needed. What areas implement what types of sensorimotor associations, precisely? What process implements these associations (e.g. thus far mainly *activation* is considered, but patterns of

⁸ There is a considerable debate about whether these models or more discrete structures better capture action control (Cooper & Shallice, 2006b, 2006a). Much of the debate surrounds whether there must be a processing or representational hierarchy *at all*. While two of us have criticized the idea of a top-down action hierarchy elsewhere (Uithol et al., 2012), we, as discussed, see no problem with positing a processing hierarchy of *some type* being present in the PFC. What is at stake is whether we describe the difference between levels in terms of increasing abstractness and context-insensitivity, or instead as differences in elements of perceptual context represented.

interaction and neural synchrony might contribute as well). What elements of the known complex anatomy in IPFC and surrounding areas implement these functions? How does the interaction of such processes produce “decisions” about what course of action to take (which are entirely attributed to discrete intentions on the folk view)? How should action understanding be framed if it is not about inferring intentions (Uithol & Paulus, 2013)? Theoretical frameworks need to be developed in order to both account for the current data and inform this type of investigation, and we have suggested that such processes should take the informational and dynamical complexity of the brain as a starting point.

7. Conclusion

Despite its importance in folk psychological explanation, philosophical accounts of action, and psychological modeling, we have argued that the notion of ‘intention’ may not be the most fruitful theoretical construct for investigation into the neural processes underlying action generation and control. There are no convincing reasons to assume the existence of such states underlying action control, and we have discussed evidence suggesting the absence of such states in the neural mechanisms that produce intentional action. We have shown that those aspects of an action that are normally thought to rely on intentions – such as temporally extended action planning and control – can be accounted for by neural processes that do not possess any of the characteristics of intentions. From this we conclude that research into intentional action is better served by studying how structures that are stable enough to control complex actions could emerge from the interaction of various control processes. Much work has to be done to expand these elements into a full-blown theory. How, for instance, are multiple types of information integrated, both across modalities and across time? How precisely are these processes shaped by the context of the action? These are difficult problems that tend to be obscured by a reliance on our intuitive understanding of intention, rooted in folk psychology. We believe that addressing them straightforwardly can lead to important progress in understanding intentional action.

Acknowledgements

The authors wish to thank Jan Derfuss, Alessandra Galli, Bernhard Hommel, Sasha Ondobaka, Michał Klincewicz, Markus Paulus, Iris van Rooij, Ivan Toni, Jason Winning and one anonymous reviewer of this journal for commenting on earlier versions of the manuscript.

References

- Anscombe, G. E. M. (1957). *Intention*. Oxford: Basil Blackwell.
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, 12(5), 193–200. <http://dx.doi.org/10.1016/j.tics.2008.02.004>.
- Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, 19(12), 2082–2099. <http://dx.doi.org/10.1162/jocn.2007.19.12.2082>.
- Barker, R. G. (Ed.). (1963). *The stream of behavior: Explorations of its structure & content*. East Norwalk, CT, US: Appleton-Century-Crofts.
- Bechtel, W. (2003). Modules, brain parts, and evolutionary psychology. In: S. J. Scher, & F. Rauscher (Eds.), *Evolutionary psychology: Alternative approaches* (pp. 221–227). Kluwer Press.
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. London: Routledge.
- Botvinick, M. M., & Plaut, D. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111(2), 395–429.
- Botvinick, M. M., & Plaut, D. (2006). Such stuff as habits are made on: A reply to Cooper and Shallice. *Psychological Review*, 113(4), 917–927.
- Bratman, M. E. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Harvard University Press.
- Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38(6), 848–863.
- Carota, F., Posada, A., Harquel, S., Delpuech, C., Bertrand, O., & Sirigu, A. (2010). Neural dynamics of the intention to speak. *Cerebral Cortex*, 20(8), 1891–1897. <http://dx.doi.org/10.1093/cercor/bhp255>.
- Christoff, K., & Gabrieli, J. D. E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology*, 28(2), 168–186.
- Christoff, K., Keramati, K., Gordon, A. M., Smith, R., & Mädlar, B. (2009). Prefrontal organization of cognitive control according to levels of abstraction. *Brain Research*, 1286(C), 94–105. <http://dx.doi.org/10.1016/j.brainres.2009.05.096>.
- Christoff, K., Ream, J. M., Geddes, L. P. T., & Gabrieli, J. D. E. (2003). Evaluating self-generated information: Anterior prefrontal contributions to human cognition. *Behavioral Neuroscience*, 117(6), 1161–1168. <http://dx.doi.org/10.1037/0735-7044.117.6.1161>.
- Churchland, P. M. (1989). *A Neurocomputational perspective*. MIT Press.
- Cisek, P. (2007). Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London, Series B*, 362(1485), 1585–1599. <http://dx.doi.org/10.1146/annurev.neuro.20.1.25>.
- Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience*, 33(1), 269–298. <http://dx.doi.org/10.1146/annurev.neuro.051508.135409>.
- Collins, A., & Koechlin, E. (2012). Reasoning, learning, and creativity: Frontal lobe function and human decision-making. *PLoS Biology*, 10(3), e1001293. <http://dx.doi.org/10.1371/journal.pbio.1001293>.
- Cooper, R., & Shallice, T. (2006a). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, 113(4), 887–916.
- Cooper, R., & Shallice, T. (2006b). Structured representations in the control of behavior cannot be so easily dismissed: A reply to Botvinick and Plaut (2006). *Psychological Review*, 113(4), 929–931.
- Courtney, S. M. (2004). Attention and cognitive control as emergent properties of information representation in working memory. *Cognitive, Affective, & Behavioral Neuroscience*, 4(4), 501–516. <http://dx.doi.org/10.3758/CABN.4.4.501>.
- Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60(23), 685–700.
- Davies, M., & Stone, T. (1995). *Folk psychology: The theory of mind debate*. Cambridge, MA: Wiley-Blackwell.
- Dietrich, E., & Markman, A. (2003). Discrete thoughts: Why cognition must use discrete representations. *Mind & Language*, 18(1), 95–119.
- Fodor, J. A. (1987). Psychosemantics: The problem of meaning in the philosophy of mind.
- Fuster, J. M. (2001). The prefrontal cortex—An update: Time is of the essence. *Neuron*, 30, 319–333.
- Fuster, J. M. (2004). Upper processing stages of the perception-action cycle. *Trends in Cognitive Sciences*, 8(4), 143–145.
- Fuster, J. M., Bauer, R. H., & Jervey, J. P. (1982). Cellular discharge in the dorsolateral prefrontal cortex of the monkey in cognitive tasks. *Experimental Neurology*, 77(3), 679–694. [http://dx.doi.org/10.1016/0014-4886\(82\)90238-2](http://dx.doi.org/10.1016/0014-4886(82)90238-2).
- Fuster, J. M., Bodner, M., & Kroger, J. K. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, 405(6784), 347–351. <http://dx.doi.org/10.1038/35012613>.
- Genovesio, A. (2006). Representation of future and previous spatial goals by separate neural populations in prefrontal cortex. *Journal of Neuroscience*, 26(27), 7305–7316. <http://dx.doi.org/10.1523/JNEUROSCI.0699-06.2006>.
- Genovesio, A., Brasted, P. J., Mitz, A. R., & Wise, S. P. (2005). Prefrontal cortex activity related to abstract response strategies. *Neuron*, 47(2), 307–320. <http://dx.doi.org/10.1016/j.neuron.2005.06.006>.
- Grafton, S. T., & Hamilton, A. F. de C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4), 590–616.
- Greenwood, J. D. (1991). The future of folk psychology: Intentionality and cognitive science.
- Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, 9(6), 290–295.
- Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews: Neuroscience*, 9(12), 934–946. <http://dx.doi.org/10.1038/nrn2497>.
- Haggard, P. (2011). Decision time for free will. *Neuron*, 69(3), 404–406. <http://dx.doi.org/10.1016/j.neuron.2011.01.028>.
- Hamilton, A. F. de C., & Grafton, S. T. (2007). The motor hierarchy: From kinematics to goals and intentions. In: P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Attention & performance 22. Sensorimotor foundations of higher cognition attention and performance* (pp. 381–408). Oxford: Oxford University Press.
- Haselager, W. F. G. (1997). *Cognitive science and folk psychology: The right frame of mind*. London, etc: Sage.
- Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C. D., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Current Biology*, 17(4), 323–328. <http://dx.doi.org/10.1016/j.cub.2006.11.072>.
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11), 494–500.
- Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, 4(11), e1000209. <http://dx.doi.org/10.1371/journal.pcbi.1000209>.
- Koechlin, E., & Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6), 229–235.

- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648), 1181–1185, <http://dx.doi.org/10.1126/science.1088545>.
- Kouneiher, F., Charron, S., & Koechlin, E. (2009). Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, 12(7), 939–945, <http://dx.doi.org/10.1038/nn.2321>.
- Lau, H. C., Rogers, R. D. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science*, 303(5661), 1208–1210, <http://dx.doi.org/10.1126/science.1090973>.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8, 529–566.
- Loh, M., & Deco, G. (2005). Cognitive flexibility and decision-making in a model of conditional visuomotor associations. *European Journal of Neuroscience*, 22(11), 2927–2936, <http://dx.doi.org/10.1111/j.1460-9568.2005.04505.x>.
- Loh, M., Pasupathy, A., Miller, E. K., & Deco, G. (2008). Neurodynamics of the prefrontal cortex during conditional visuomotor associations. *Journal of Cognitive Neuroscience*, 20(3), 421–431.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31(5), 838–850.
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
- Momennejad, I., & Haynes, J.-D. (2012). Human anterior prefrontal cortex encodes the 'what' and 'when' of future intentions. *NeuroImage*, 61(1), 139–148, <http://dx.doi.org/10.1016/j.neuroimage.2012.02.079>.
- Moore, J. W., Wegner, D. M., & Haggard, P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, 18(4), 1056–1064, <http://dx.doi.org/10.1016/j.concog.2009.05.004>.
- Mushiaki, H., Saito, N., Sakamoto, K., Itoyama, Y., & Tanji, J. (2006). Activity in the lateral prefrontal cortex reflects multiple steps of future events in action plans. *Neuron*, 50(4), 631–641, <http://dx.doi.org/10.1016/j.neuron.2006.03.045>.
- Ondobaka, S., Lange, F. P., Newman-Norlund, R. D., Wiemers, M., & Bekkering, H. (2012). Interplay between action and movement intentions during social interaction. *Psychological*.
- Ongür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex*, 10(3), 206–219.
- Pacherie, E. (2000). The content of intentions. *Mind & Language*, 15(4), 400–432.
- Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107(1), 179–217.
- Pacherie, E. (2011). Nonconceptual representations for action and the limits of intentional control. *Social Psychology*, 42(1), 67–73, <http://dx.doi.org/10.1027/1864-9335/a000044>.
- Pacherie, E., & Haggard, P. (2010). What are intentions?. In: W. Sinnott-Armstrong, & L. Nadel (Eds.), *Conscious will and responsibility. A tribute to Benjamin Libet* (pp. 70–84). Oxford: Oxford University Press.
- Passingham, R. E., Toni, I., & Rushworth, M. F. S. (2000). Specialisation within the prefrontal cortex: The ventral prefrontal cortex and associative learning. *Experimental Brain Research*, 133(1), 103–113.
- Petrides, M. (2005). The rostral-caudal axis of cognitive control within the lateral frontal cortex. In: S. Dehaene, J.-R. Duhamel, M. D. Hauser, & G. Rizzolatti (Eds.), *From monkey brain to human brain. A Fyssen foundation symposium* (pp. 293–314). Cambridge, MA: MIT Press.
- Pezzulo, G., & Castelfranchi, C. (2009). Intentional action: From anticipation to goal-directed behavior. *Psychological Research-Psychologische Forschung*, 73(4), 437–440.
- Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, 3(1), 85–90, <http://dx.doi.org/10.1038/71156>.
- Ramnani, N., & Owen, A. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews: Neuroscience*, 5(3), 184–194.
- Ramsey, W., Stich, S., & Garon, J. (1999). Connectionism, eliminativism and the future of folk psychology. In: J. D. Greenwood (Ed.), *The future of folk psychology: Intentionality and cognitive science* (pp. 93–119). Cambridge: Cambridge University Press.
- Ranganath, C., Johnson, M. K., & D'Esposito, M. (2000). Left anterior prefrontal activation increases with demands to recall specific perceptual information. *Journal of Neuroscience*, 20(22), RC108.
- Rogers, R. D. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology General*, 124(2), 207.
- Rushworth, M. F. S. (2008). Intention, choice, and the medial frontal cortex. *Annals of the New York Academy of Sciences*, 1124(1), 181–207, <http://dx.doi.org/10.1196/annals.1440.014>.
- Sakai, K. (2008). Task set and prefrontal cortex. *Annual Review of Neuroscience*, 31, 219–245, <http://dx.doi.org/10.1146/annurev.neuro.31.060407.125642>.
- Schurger, A. A., Sitt, J. D. J., & Dehaene, S. S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences of the United States of America*, 109(42), E2904–E2913, <http://dx.doi.org/10.1073/pnas.1210467109>.
- Searle, J. (1983). *Intentionality, an essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Simons, J., Schölvinck, M., Gilbert, S. J., Frith, C. D., & Burgess, P. W. (2006). Differential components of prospective memory? Evidence from fMRI. *Neuropsychologia*, 44(8), 1388–1397, <http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.005>.
- Soon, C. S., Brass, M., Heinze, H.-J., & Haynes, J.-D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11(5), 543–545, <http://dx.doi.org/10.1038/nn.2112>.
- Soon, C. S., Hanxi, He, A., Bode, S., & Haynes, J.-D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences*, 1–6, <http://dx.doi.org/10.1073/pnas.1212218110>.
- Stich, S. P. (1983). *From folk psychology to cognitive science: The case against belief*. Cambridge, MA: The MIT Press.
- Uithol, S., & Paulus, M. (2013). What do infants understand of others' action? A theoretical account of early social cognition. *Psychological Research*. <http://dx.doi.org/10.1007/s00426-013-0519-3>.
- Uithol, S., van Rooij, I., Bekkering, H., & Haselager, W. F. G. (2012). Hierarchies in action and motor control. *Journal of Cognitive Neuroscience*, 24(5), 1077–1086, http://dx.doi.org/10.1162/jocn_a_00204.
- Wegner, D. M. (2003). *The illusion of conscious will*. Cambridge, MA: The MIT Press.
- Wenke, D., Fleming, S. M., & Haggard, P. (2010). Subliminal priming of actions influences sense of control over effects of action. *Cognition*, 115(1), 26–38, <http://dx.doi.org/10.1016/j.cognition.2009.10.016>.
- Yamashita, Y., & Tani, J. (2008). Emergence of functional hierarchy in a multiple timescale neural network Model: A humanoid robot experiment. *PLoS Computational Biology*, 4(11), 1–18.