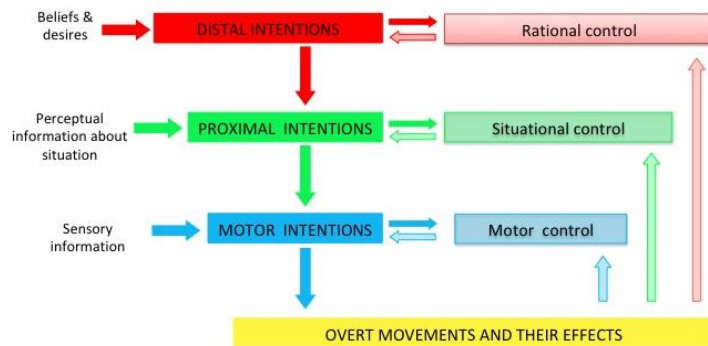## Intentions: The dynamic hierarchical model revisited

Myrto Mylopoulos (Carleton University) & Elisabeth Pacherie (Institut Jean Nicod)

Abstract: Ten years ago, one of us proposed a dynamic hierarchical model of intentions, the DPM model, that brought together philosophical work on intentions and empirical work on motor representations and motor control (Pacherie 2008). The model distinguished among Distal intentions, Proximal intentions and Motor intentions operating at different levels of action control. This model specified the representational and functional profiles of each type of intention, as well as their local and global dynamics, and the ways in which they interact. A core insight of the model was that action control is the result of integrated, coordinated activity across these levels of intention. Since the proposal of the model, empirical and theoretical work in philosophy and cognitive science has emerged that would seem to support and expand on this central insight. In particular, an updated understanding of the nature of sensorimotor processing and motor representations, as well as of how the different levels of intention and control interface and interact, allows for the further specification and precisification of the original DPM model.

Graphical/Visual Abstract and Caption

The DPM model: a dynamic hierarchical model of intentions with its three main levels of action specification and action control.

Introduction

About a decade ago, one of us proposed a hierarchical model of intentions, the DPM model, that made a distinction between Distal intentions, Proximal intentions and Motor intentions (Pacherie 2006, 2007, 2008). The development of this model was motivated in part by dissatisfaction with existing answers to one of the most central questions in the philosophy of action, namely: what makes actions actions? Or, as Wittgenstein put it: "When I raise my arm, my arm goes up. And the problem arises: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?" (Wittgenstein 1953; 1, § 621). The standard answer provided by proponents of the causal theory of action was that when you subtract the movement from an action, you are left with the mental states that are the causal antecedents of the

movement. These philosophers did not agree, however, on what these mental states were supposed to be. In addition, the accounts proposed generally failed to capture the most proximal mental antecedents of action, leaving them prey, among other things, to problems of deviant causation. A second, positive, motivation for offering this model was the conviction that empirical work in what was then the emerging field of motor cognition was key to solving these difficulties. In particular, the motor representations investigated by neuroscientists appeared as promising candidates for the role of most proximal causal antecedents of action. Finally, another important motivation for the model was the naturalist conviction that philosophical progress in many areas depends on intimate engagement with empirical science and the corresponding desire to encourage dialogue among philosophers and scientists working on action.[1]

Things have changed. While it would certainly be an exaggeration to say that philosophers of action have converted en masse to naturalism, a growing number are now willing to take into account empirical findings about actions and their cognitive underpinnings (for recent work see, e.g., Christensen et al. 2016; Ferretti 2016; Fridland, 2017; Grünbaum 2017, Nanay, 2013a, Shepherd 2016, Wu 2016, Wong 2015). Talk of motor intentionality and motor control doesn't ring as foreign to philosophical ears as it once did and the notion of "motor representation" has come under increased philosophical scrutiny. During the same period, the cognitive neuroscience of action has become a well established field yielding a wealth of new empirical findings and theoretical models, highly relevant to thinking about motor intentionality and representational and control hierarchies.

Our first aim, in section 2, is to summarize the key concepts in the DPM model and to offer clarifications regarding certain aspects of the original model. Our second aim is to refine and update this model in the light of new philosophical and empirical work. In section 3, we revisit motor representations. In section 4, we discuss three sets of issues pertaining to coordination across levels of intention: to what extent do higher-level intentions determine the content of lower-level intentions? How do they communicate with each other? What are the respective contributions to action control of intentions at different levels of the hierarchy?

THE DPM MODEL

The DPM model distinguishes three main stages in the process of action specification, each corresponding to a different level or layer of intention, and intentions at each level having a distinctive role to play in the guidance and monitoring of the action. It thus proposes a threefold distinction among distal, proximal, and motor intentions (D-intentions, P-intentions and M-intentions for short). The model also offers a characterization of the two levels of dynamics involved in the unfolding of intentions. The local dynamics specific to each level of intention can be divided into two phases: an upstream phase culminating in the formation of the intention and a downstream phase concerned with the monitoring and control role of the intention once formed. The global dynamics refers to the transitions from one level of intention to the next, where these transitions involve the construction of progressively more detailed representations of the action to be performed.

A number of philosophers had previously developed dual-intention theories of action, variously distinguishing between prior intentions and intentions-in-action (Searle 1983), future-directed and present-directed intentions (Bratman 1987), prospective and immediate intentions (Brand 1984), or distal and proximal intentions (Mele 1992). While the DPM model builds on some of the insights of these dual-intention theories, it departs from them in two main ways. First, and most obviously, it introduces motor intentions and thus a threefold rather than a twofold distinction among intentions. Second, rather than assuming that the role of distal

intentions is over once proximal intentions have been formed, the DPM model takes it that intentions at each of the three levels have a distinctive role to play in the control and monitoring of ongoing actions.

In the DPM model, the distinction between D-intentions, P-Intentions and M-intentions is motivated by an analysis of their different and complementary functional roles, of the different types of contents they involve and of their respective temporal scales.

D-intentions are in certain respects close kins to Bratman's future-directed intentions (Bratman 1987). D-intentions function as terminators of practical reasoning about ends, prompters of practical reasoning about means and plans, and intra- and interpersonal coordinators and are subject to norms of practical rationality. Once one has settled on a goal, one should form a plan for achieving this goal that is internally consistent, consistent with one's beliefs about the world (external consistency), and consistent with other intentions and projects one has already settled on (global consistency).

It bears emphasis that the DPM model is a hierarchical model not just in the sense that it takes D-intentions, P-intentions and M-intentions to be hierarchically organized, but also in that it allows for a hierarchical organization within layers. Thus once one has settled on a goal, say formed the D-intention to attend a certain conference, planning may involve the formation of more specific D-intentions, like arranging for travel, booking accommodation, paying conference fees, etc., corresponding to subgoals or sub-steps towards this overarching goal.

The practical rationality constraints that bear on D-intentions require the presence of a network of inferential relations among intentions, beliefs, and desires to insure internal, external and global consistency. It is generally admitted that concepts are the inferentially relevant constituents of intentional states and that their sharing a common conceptual representational format is what makes possible a form of global consistency, at the personal level, of our desires, beliefs, intentions and other propositional attitudes. If we accept this common view, it follows that in order to satisfy the rationality constraints they are subject to, D-intentions must have a propositional format and conceptual content.

In addition, to the extent that D-intentions can concern future actions to be performed in situations on which the agent has as yet very little information, their contents will often be not just conceptual but also largely descriptive. Note though, that this doesn't mean that the contents of D-intentions must remain purely descriptive. The content of D-intentions can involve indexical elements if one already finds oneself in the context in which one plans to act in the future (e.g., I will eat this piece of chocolate in front of me as soon as I have finished writing this section). In addition, in some cases, an agent may be in a position to imagine the future situation of action and thus their action in great detail. This may allow the agent to form what Gollwitzer (1999) calls "implementation intentions". In contrast to goal intentions that only specify what one will do, implementation intentions are if-then plans that also specify the when, where and how of what one will do and tie the goal-directed behavior to anticipated situational cues. There is ample empirical evidence that implementation intentions enhance the likelihood of goal achievement compared to goal-intentions (see Gollwitzer and Sheeran (2006) and Prestwich et al. (2015) for reviews). According to Gollwitzer, implementation intentions achieve their success by delegating the control of goal-directed behavior to anticipated situational cues, which, when actually encountered, are easily detected and elicit goal-directed behavior automatically. As Pacherie and Haggard (2010) point out, this early planning strategy may be better suited to some situations than others. If the agent's anticipations are not correct, the situational cues on which action initiation depends may fail to materialize. Or, worse perhaps, the cues may be present and automatically trigger the action when other unanticipated and unattended aspects of the situation

make it unadvisable to pursue as planned. Conversely, late planning allows for more flexibility but runs the risk of unpreparedness when the time of acting comes.

Finally, in contrast to Bratman's view, for whom when the time to act comes future-directed intentions give way to present-directed intentions that are solely in charge of action control, the DPM model assigns a specific control function to D-intentions. They are in charge of the rational control of the action, that is of insuring that the way the action unfolds – in particular, unanticipated side effects or adjustments of the action plan in the face of unexpected difficulties –doesn't flout the reasons the agent had for their action in the first place or violate norms of practical rationality.

A P-intention often, though not always, inherits an action plan from a D-intention. Its task is then to anchor this plan in the situation of action. The temporal anchoring, the decision to start acting now is but one aspect of this process. Once the agent has established a perceptual information-link to the situation of action, she must insure that the action plan is implemented in that situation. This means that she must generate an indexical representation of the action that fits the specification inherited from the D-intention while anchoring it to the situation at hand. The formation of a P-intention depends on the integration of conceptual information about intended action inherited from the D-intention with perceptual information about the current situation and memory information about one's motor repertoire to yield a more definite representation of the action to be performed. Once formed P-intentions are in charge of the situational monitoring and control of the action as it unfolds. This control is more local than the rational control exerted by D-intentions and is concerned with the immediate goal and the situation as currently perceived.

The main novelty of the DPM model is perhaps the introduction of motor representations. This introduction was motivated by the thought that philosophical accounts of action in terms of (traditional) intentions were incomplete, leaving an explanatory gap. This gap had not gone unnoticed, but philosophers were at a loss how to fill it. The notorious problem of deviant causation, already pointed out by Davidson (1963), illustrates this gap. Suppose, to borrow an example from Davis (1994), that John intends to get down on his knees to propose marriage, but that arriving in front of his beloved, he is so overcome with emotion that he suddenly feels weak and sinks to his knees. Clearly, John's sinking to his knees is not an action even though it is caused by his intention to get down on his knees. The existence of a causal relation between an intention representing a certain outcome and a corresponding outcome doesn't guarantee that the production of this outcome is a purposive action. The challenge is then to specify what causal connection must hold between the intention and the resultant behavior for the latter to qualify as an action. In answer to that challenge, the DPM model proposes that in many instances of deviant causation, the problem can be diagnosed as arising from the fact that the causal chain leading from intention to behavior fails to include motor intentions. Thus, causing "in the right way" may be a matter of intentions causing behavior via the instantiation of appropriate motor intentions.

M-intentions involve what neuroscientists call motor representations. We will consider in the next section some recent debates about the nature of motor representations. In the original DPM model, it was proposed that motor representations encode action goals together with the motoric means for achieving them and do so in a motoric format directly suitable to action execution. The construction of motor representations depends in part on processing by the vision for action system (Milner and Goodale 1995), whose role is to extract from visual stimuli information about the action-relevant properties of objects and situations, information that is then used to compute motor commands appropriate to the goal pursued. Motor representations are not just representations of movements, rather they represent goal directed movements and

the way they encode movements is normally sensitive not just to the immediate motor goal but to the hierarchy of motor goals in which they are embedded. For instance, the same cup will be grasped in different ways depending on whether one wants to bring it to one's lips or to turn it upside down. Thus, motor representations appear to be themselves hierarchically organized, with a decomposition of more general into more specific motor goals (Jeannerod, 2006, Grafton & Hamilton 2007, Van El et al. 2014).

According to the DPM Model, M-intentions are also responsible for a distinctive form of monitoring and control of an ongoing action. They are responsible for the precision and smoothness of action execution and motor control typically operates automatically and at a very fine time-scale.

Why talk of M-intentions rather than simply motor representations? M-intentions have the functional role of bringing about the actions they represent, whereas motor representations need not have such a functional role. For instance, motor representations are also formed when we observe someone else acting or when we imagine acting (Jeannerod 1995; Filimon et al. 2007). Thus, talk of M-intentions emphasizes the role of such states in action production and the fact that M-intentions share certain core functional characteristics of D-intentions and P-intentions. They represent goals and means to those goals. They have, in Searle's terminology, a world-to-mind direction of fit: in achieving success the world is altered to fit the content of M-intentions. Finally, they have, in Searle's terminology again, a mind-to-world direction of causation: they cause the realization of what they represent.

Brozzo (2017) has objected that M-intentions fail to meet three constraints that genuine intentions should satisfy, namely conscious accessibility, integration with other propositional attitudes and strong consistency, where conscious accessibility and integration are taken as preconditions of strong consistency. Brozzo is right that according to the DPM Model, we may only have partial access to the contents of M-intentions and that there are restrictions on the kind of information directly available to the motor system. To that extent, the use of the term *intention* for M-intention departs in some respects from traditional usage. The question then is whether there is more at stake here than mere terminological orthodoxy.[2] M-intentions' failure to meet the constraints highlighted by Brozzo would be problematic if it led to systematic inconsistencies between an agent's D-intentions or P-intentions and the behavior for which M-intentions are immediately responsible. However, the architecture of the DPM model precludes such systematic inconsistencies. On the one hand, what M-intentions are formed and what contents they have is constrained by P-intentions and their contents, so that initially at least M-intentions will not be inconsistent with higher-level intentions. On the other hand, the existence of three levels of monitoring and control normally insures that emerging discrepancies during action execution are detected and reduced. Thus, while M-intentions may not, as Brozzo points out, directly meet strong consistency constraints, the architecture in which they are embedded insures that they meet them indirectly.

MOTOR REPRESENTATIONS REVISITED

As pointed out by Ferretti (2016), there is as of yet no consensus in the literature on the contents of motor representations, on their conscious accessibility and on their sources. Some hold that motor representations encode mainly action goals or action outcomes (Butterfill and Sinigaglia 2014), others that they represent action properties, as distinct from either goals or movements (Nanay 2013a), and yet others that they represent action goals together with motor commands specifying the movements to be executed to achieve those goals (Jacob & Jeannerod 2003, Jeannerod 2006, Pacherie 2008). In addition, there is also disagreement as to whether motor representations are entirely unconscious (Milner and Goodale 1995), or at least partly

accessible to consciousness, and under what conditions (Brozzo 2017, Gallese 2007, Nanay 2013a, Pacherie 2008). The conscious accessibility issue is related, in turn, to the source issue: are motor representations solely the product of dorsal processing in the visual stream? Or do both the dorsal and the ventral streams contribute to the formation of motor representations?

Thinking about motor representations has been strongly and perhaps overly influenced by Milner & Goodale's dual model of vision (Milner & Goodale 1995, 2006) and for a long time has relied almost exclusively on the study of reach and grasp actions.[3] In recent years however, aspects of this model have been challenged. Greater attention has been given to other kinds of action, in particular actions involving tool use, and there has been greater emphasis on the idea that motor representations must integrate information from different sources and are themselves hierarchically organized along a gradient of abstraction (Grafton & Hamilton 2007; Hamilton & Grafton 2007). Together these developments suggest ways of resolving some of the tensions pointed out by Ferretti (2016). In what follows, we survey some of these developments, using Milner and Goodale's model as our starting point.

Milner & Goodale's dual model of vision (Milner & Goodale 1995, 2006) interprets the anatomical division between a dorsal visual pathway and a ventral visual pathway as corresponding to a functional division, the dorsal pathway being responsible for the visual guidance of actions directed at objects in the environment (vision-for-action) and the ventral pathway for the construction of visual percepts and conscious object perception (vision-for-perception). According to this model, vision-for-action and vision-for-perception differ in four main ways: 1. Vision for action encodes spatial properties using egocentric frames of reference, while vision for perception encodes them allocentrically, which is part of the explanation for why vision-for-action is largely unaffected by contextual illusions. 2. Vision for action is optimized for action guidance and thus builds transient representations of rapidly changing spatial properties and relations, whereas vision-for-perception constructs sustained representations. 3. Vision-for-action is informationally encapsulated and largely insensitive to ventral stream processing. 4. Only the products of vision-for-perception reach visual awareness, vision-for-action remaining entirely unconscious.

Early thinking about motor representations was strongly influenced by this model. Motor representations were considered to be the product of dorsal processing and thus thought to inherit the properties attributed to the vision-for-action system: transience, egocentric spatial coding, informational encapsulation and inaccessibility to consciousness.

In recent years, however, a growing body of empirical literature has challenged Milner and Goodale's dichotomous characterization of the two visual systems, suggesting that their functional independence is not as strong as envisioned on their model. For instance, there is neurophysiological and neuroimaging evidence for allocentric spatial coding in the dorsal stream (Adam et al. 2016, Kravitz et al. 2011). And many recent studies have also reported that visual illusions can affect actions, although this effect is typically weaker than their effect on perceptual judgments (see, e.g., Bruno, Bernardis, & Gentilucci, 2008; Bruno & Franz, 2009; Franz & Gegenfurtner, 2008, Schenk & McIntosh 2010, Kopiske et al. 2016).

Substantial evidence has also accrued that the anatomical and functional organization of the dorsal stream is more complex than initially thought. Thus, Rizzolatti and Matelli (2003) have described two anatomically segregated subcircuits of the dorsal stream, a dorso-dorsal pathway projecting from visual areas V3 to V6, via areas in the superior parietal lobule, to dorsal pre-motor areas and a ventro-dorsal pathway, projecting from medial superior temporal area (MT/MST) ventral premotor cortex via the inferior parietal lobule (IPL).[4] Based on an analysis

of the different deficits associated with lesions in the dorso-dorsal and the ventro-dorsal streams, it has been proposed that the two streams subserve two distinctive action systems (Binkofski and Buxbaum 2013; Pisella et al. 2006). The dorso-dorsal "grasp" system processes structural characteristics of currently-viewed objects and is particularly relevant for the online control of prehensile actions. The mostly left-lateralized ventro-dorsal "use" system, in contrast, processes sensorimotor information based upon longer-term object use representations, and is critical for skilled, functional object use. It is thought to store sensorimotor "engrams" or "schemas" specifying the invariant characteristics of familiar skilled actions (Buxbaum et al. 2007; Niessen et al. 2014; Vry et al. 2015). (For more on motor schemas, see section 4.)

In addition, neuroanatomical and neurofunctional studies have uncovered many connections between the dorsal substreams and the ventral stream, indicating that these streams are able to communicate with each other in a bidirectional way at various processing stages and suggesting that the ventro-dorsal substream may constitute an important interface between the ventral and the dorso-dorsal streams of visual information processing. Similarly, brain imaging studies have shown that the dorsal and ventral streams are often jointly implicated in situations involving delayed or pantomimed grasping as well as in tool use, where semantic knowledge needs to be accessed to allow for the selection of the most appropriate grasp (for reviews, see Cloutman 2013; Grafton 2010). The existence of both anatomical and functional interconnectivity between the dorsal and ventral streams suggests the existence of some form of informational exchange between these streams.

In addition, the extent to which visually guided action is cognitively penetrable and the ways in which it may be influenced by semantic information are matters of intense current debate (Briscoe & Schwenkler 2015; Brogaard, 2011; Kozuch, 2015; Mahon & Wu, 2015; Nanay, 2013; Shepherd, 2015; Wu, 2014; Zipoli Caiani & Ferretti, 2017). As pointed out by Mahon and Wu (2015), there are "three places that semantic information can influence dorsal stream computations: (a) at the level of input to the dorsal stream; (b) at the level of output from the dorsal stream; and (c) at the level of internal computations within the dorsal stream" (2015: 211). On their view, only the third kind of influence would count as cognitive penetration. Milner and Goodale acknowledged that vision-for-perception interacts with vision-for-action by identifying potential goal objects, but they drew a sharp distinction between a planning stage where the two systems interact and programming and on-line control stages which depend solely on the vision-for-action system. The only points at which the two systems could make contact were, according to them, common retinotopic maps. Thus, on their model the only way the semantic information from vision-for-perception could influence vision-for-action would be by "highlighting" particular regions in retinotopic maps for further processing by the vision-for-action system and thus by influencing the input to dorsal stream computations.

Alternatively, one may consider that the detection of action possibilities and the selection of an action plan constitute two successive steps. The dorsal stream would automatically detect the action possibilities and motor plans offered by a particular situation and the objects therein and semantic information about the agent's goal would then be used to select one among all these action possibilities and motor plans (for a detailed critical discussion of this approach, see Zipoli Caiani & Ferretti 2017). This would correspond to the second possibility delineated by Mahon & Wu (2015). However, having two segregated detection and selection stages wouldn't be very efficient from a computational point of view, given the large number of action possibilities a situation usually affords and the interference problems that their simultaneous activation would be likely to create. Thus, it might make better computational sense if semantic information could influence not just inputs or outputs but the actual computations carried out by the dorsal stream.

There are now several lines of evidence that semantic information can bias the computations internal to the dorsal stream. For instance, Jeannerod, Decety, and Michel (1994), reported the case of a patient with a lesion in the dorsal stream whose grasping movements were much less inaccurate when grasping familiar objects than when grasping novel objects, suggesting that knowledge of object identity could compensate for impaired online visuomotor processing. The influence of semantic knowledge of object identity on the programming of grasping has been further demonstrated in a more recent study by McIntosh & Lashley (2008). In their study, they used two brands of matchboxes very popular in the UK. They exploited the fact that the boxes in brand A were 25% larger than the boxes in brand B and made replicas of the boxes, scaling down brand A boxes to the size of brand B boxes and vice versa. They found that brand identity influenced grip size, with participants using larger grip size when grasping the scaled-down replica of the brand A box than when grasping the normal brand B box of the same size (for discussion, see Nanay 2013).

However, the main way in which semantic information influences dorso-dorsal processing is probably through the mediation of the ventro-dorsal stream. In particular, semantic information about the functional properties of objects stored in the inferior parietal lobule constrains the way their structural properties are processed.

These findings also constitute a challenge to Milner and Goodale's view that the representations constructed by the vision-for-action system are inaccessible to consciousness. Their supposed transience was one of the main reasons advanced for denying their conscious accessibility. However, even if representations computed by the dorso-dorsal "grasp" system may indeed decay too rapidly to reach consciousness, empirical evidence suggests that the representations computed by the ventro-dorsal "use" system are maintained over longer periods of time and thus that their temporal characteristics would be compatible with accessibility to consciousness. This point relates to the familiar debate on the extent to which the visual guidance of behaviour provided by the dorsal stream is akin to "zombie action" (see, e.g., Mole 2009 ; Wu 2013 ; Shepherd 2015, 2016). It may be that the low-level, online computations of the dorso-dorsal stream, e.g, those specifying object size for the purposes of programming grip aperture during reaching movements, are indeed inaccessible to consciousness. Yet, action guidance is not necessarily entirely "zombie-like" insofar as consciously accessible and more stable representations in the ventro-dorsal "use" system are involved in computing motor schemas corresponding to certain action targets and types.

As for our conscious access to the M-intentions produced by dorsal vision computations, again there seems to be evidence that it is blocked in the context of online action guidance. This much is suggested both by everyday experience and experimental results. To take the latter, consider a now-classic experiment by Castiello, Paulignan, and Jeannerod (1991), in which participants were asked to reach for targets that suddenly changed position midway through their movements, often during a saccade. Participants had no trouble performing online adjustments of their movements in response to the target shift, however they were not aware until almost half a second after the object had been displaced that it had moved, nor were they aware of the automatic adjustments they made. (See also Fourneret & Jeannerod 1998; Chen & Saunders 2016).[5]

Despite the seeming inaccessibility of M-intentions involved in online action guidance, some have appealed to motor imagery to argue that such representations might nonetheless be sometimes consciously available (Jeannerod 1994, 2006; Brozzo 2017; Shepherd 2017). Motor imagery occurs when an agent imagines herself executing an action from a first-person perspective. It has been discovered that fine-grained parameters of the imagined action, such as temporal and spatial configuration, correspond to those of the executed action. From this

observation, some have concluded that an agent consciously accesses the M-intentions activated during such imagery (e.g., Brozzo 2017, p. 243). The difficulty here is that it is hard to rule out the alternative possibility that what an agent is aware of during motor imagery is restricted to kinaesthetic experiences, rather than the motor representations that give rise to them.

At the very least, as Shepherd (2017) argues, there does seem to be some indication that motor representations affect the content and subjective character of experiences of acting, and in that way might be indirectly accessed. Shepherd points to evidence that in the active version of the Rubber Hand Illusion, in which participants perform finger movements corresponding to movements of the rubber hand, proprioceptive "drift" affects the whole hand, while in the passive condition, which merely involves stroking, it is restricted to the relevant finger (Kalckert & Ehrsson 2014). The explanation offered is that spatial information specified by motor representations in the primary motor cortex has a receptive field that includes the entire hand and thus gets integrated with spatial information provided by representations in the primary somatosensory cortex in the active, but not the passive, condition. This results in experiential content that is jointly influenced by motor and somatosensory representations, thus allowing for indirect access to the contents of our M-intentions.

The relaxation of strictures on motor representations – transience, egocentric spatial coding, informational encapsulation and inaccessibility to consciousness – supported by recent work is welcome news for the DPM Model. Thus, in its original version it claimed that the way in which motor representations encode movements is normally sensitive not just to immediate motor goals but to the hierarchy of motor goals in which they are embedded, it did not offer a clear way of fitting this claim with the idea that motor representations are the product of computations by an encapsulated vision-for-action system. The more complex picture that has emerged of vision-for-action and of the interactions between dorso-dorsal, ventro-dorsal and ventral pathways makes for an easier fit. For instance, it suggests a possible division of labor between the ventro-dorsal stream storing motor schemas or engrams and thus more involved in the representation of goals and their organization into subgoals and the dorso-dorsal stream more involved in the parametrization of these motor schemas based on current constraints imposed by the body and environment. Such a division of labor would be consistent with the claim by Butterfill and Sinigaglia (2014) that certain motor representations encode mainly action goals. In addition, as we will see in the next section, these developments make the problem of coordination across levels of intention appear less daunting.

COORDINATION ACROSS THE LEVELS: DETERMINATION, INTERFACING, AND DIVISION OF LABOUR

A general upshot of the DPM model is that D-intentions, P-intentions, and M-intentions coordinate and integrate with one another in the service of action control. This immediately raises a number of questions. First, there is the question of the extent to which intentions at higher levels in the hierarchy determine the content of those at lower levels. Call this the *Determination Question*. Next, there is the question of how it is that propositionally structured intentions at higher levels in the hierarchy communicate with the motorically formatted representations at lower levels. Call this the *Interface Question*. Finally, there is the question of how we should understand the relative contribution to overall action control from cognitive control mechanisms, located at the levels D-intentions and P-intentions, and automatic motor control mechanisms, located at the level of M-intentions. Call this the *Division of Labour Question*. We take each of these questions in turn.

Determination vs. Biasing

Recall that on the DPM model, each level of intention—D-intention, P-intention, and M-intention—possesses both local and global dynamics. Local dynamics are restricted to a given level, while global dynamics involve transitions from one level to the next. These transitions, in turn, result in the formation of increasingly more detailed goal representations.

How do the transitions between higher and lower levels of intention yield the specific contents that they do? Following Burnston (2017), we can distinguish between two ways of answering this question. On the *determination view*, the transition between levels of intention goes like this: the tokening of an intention X at a given level with content C1 results, by way of a unidirectional process of content-causation, in the tokening at the next level of an intention Y, with content C2. On this view, the tokening of X with its particular content C1, along with background conditions, is sufficient for the tokening of Y with its particular content C2. On the *biasing view*, by contrast, the tokening of an intention X, biases the content of Y, but does not determine it. Rather, the role of the higher-level intention is to "potentiate" a range of possible contents at the next lowest level. The specific content that is formed is heavily influenced in addition by context and perceptual processing.

Determination views seem implausible. Against such views, Burnston argues that they fail to account for the diversity and specificity of sensorimotor representations corresponding to a token intention. Suppose one intends to pick up the drinking cup in front of them. In order for this intention to be implemented, a number of movement parameters must be specified, including what type of hand grip to use, how much force, and the direction and angle of movement, all depending on the size, shape, location, and composition of the cup. There is a range of values that these parameters might take. These values are not "filled in" by the intention, but by downstream processing that results in the appropriate sensorimotor representations. Thus, intentions and background conditions are not alone sufficient for determining the content of sensorimotor representations. Applied to the three-fold distinction between D-intentions, P-intentions, and M-intentions, the lesson is that P-intentions *bias* but do not *determine* the content of M-intentions.[6]

How might this biasing work? An emerging consensus is that attention plays an important role. The precise nature of this role is an area that is ripe for future research. One promising account is offered by Wu (2016), a proponent of the biasing approach (see esp. p. 107). In line with standard empirical operationalisations of attention, Wu understands it to be an individual-level process by which perceptual information is selected for the purpose of a task. And he views intentions as states (vs. events) that modulate attention in such a way as to bias the coupling of goal-appropriate perceptual inputs with goal-appropriate behavioural responses. It is this modulation of attention by way of intention that allows the correct perceptual targets within an action context to be isolated and appropriately acted upon. Not only this, but the perceptual information that attention selects is then used in the programming of the relevant M-intention (p.113).

Wu's (2016) proposal dovetails nicely with empirical work on the function of visual attention, which suggests that there are two segregated, though interacting, neural systems implicated in the control of visual attention for action (for reviews, see Corbetta & Shulman 2002; Corbetta, Patel, & Shulman 2008). The first system, located in the dorsal posterior parietal and frontal cortex, is primarily implicated in the perception-action coupling that Wu describes. This system is responsible for two specific functions related to such coupling, both guided by the agent's intention: (i) preparation of stimulus-selection in accordance with expected attributes of

perceptual targets (e.g., location, motion, colour), and (ii) preparation of corresponding response-selection in accordance with the relevant task.

The second system, primarily involving the temporoparietal junction (TPJ) and ventral frontal cortex, differs from the first system in that it is not activated by way of prior expectations. It has the function of detecting salient targets relevant to the present goal, often at unattended locations, and directing and orienting attention to those stimuli. This frequently takes the form of a re-orientation of attention from a cued location to the surprise location of an unexpectedly appearing target. Intention thus also modulates attentional mechanisms by priming them to be sensitive to certain classes of stimuli, even if these are not present or even expected at the time of goal formation.

The idea of perception-action coupling is also central to the Theory of Event Coding (Hommel et al. 2001), which offers another way of fleshing out, in general terms, how biasing works (Burnston 2017). This theory treats sensorimotor representations as bindings of certain values along perceptual and motoric dimensions, some of which can be formed in an automatic and associative manner. For instance, Burnston (2017) discusses evidence that perceiving a grippable object (e.g., a hammer) automatically activates the standard grip associated with that object, since being asked to grip it in a non-standard way results in slower reaction times (Tucker and Ellis 2001). Biasing accounts thus have at their disposal rich interactions among intentions, attention, and coupling processes to explain how transitions from one level of intention to the next can result in further specified contents that outstrip in detail the contents provided at higher levels.

Interfacing

Regardless of exactly how biasing works, all can agree that successful action requires that joint coordination takes place *in some way* between intentions and the motor system. But here one is faced with the following problem: How is it that D- and P-intentions, with their propositional format, are able to successfully coordinate with M-intentions, given their non-propositional, motoric format? Following Butterfill & Sinigaglia (2014) we call this the *Interface Problem*.[1]

At first glance, it would seem that what is needed to solve the interface problem is some sort of translation process taking place between D- and P-intentions and M-intentions. Take the analogy of communication between two individuals who speak different languages. In order to successfully communicate with one another, they must translate the content of their utterances into a common language that they both understand. But Butterfill & Sinigaglia (2014) protest that nothing at all is known about any such mysterious translation process when it comes to D- and P-intentions and M-intentions, and so we must come up with a different type of solution.

---

[1] We note here that the interface problem is also a problem for recent hybrid accounts of skilled action that attempt to explain the intelligence of skill not merely by appeal to propositional knowledge or trained up motor dispositions, but rather by appeal to how control processes at both cognitive and motoric levels integrate and coordinate, with neither level being privileged as *the* locus of control. (see, e.g., Fridland 2014, 2016; Levy 2017; Christensen et al. 2016; Papineau 2015; Shepherd 2017).

Here we review three proposed solutions, including that of Butterfill & Sinigaglia (2014): (i) the deferral solution (Butterfill & Sinigaglia 2014), (ii) the motor schema solution (Mylopoulos & Pacherie 2017), and (iii) the dual-content solution (Shepherd 2017).

The deferral solution appeals to the idea that P-intentions have demonstrative content, which allows them to refer to actions by way of deferring to motor representations. Butterfill & Sinigaglia (2014) illustrate their proposal by way of an analogy with a visual map and spoken directions, the former taking pictorial format and the latter taking propositional format. They argue that one can refer to a specific route depicted by a line on the map by way of saying, "Follow *that* route!" and pointing to the line. In this case, 'that' is a demonstrative referring to the route by deferring to the line on the map. Similarly, P-intentions with the content <Do that!> can refer to action outcomes by way of deferring to M-intentions that specify those same outcomes.

While the deferral solution may be tempting (for recent endorsements see Levy 2017 and Brozzo 2017), one significant worry one might have is that it presupposes the very translation process it aims to avoid. For just as one must know *which line* on the map to defer to in order to refer to the correct route, P-intentions must select the appropriate M-intentions to defer to in order to refer to the correct action outcome. The deferral solution is silent as to how this works, and so the interface problem still looms.

Another proposed solution to the interface problem, by Mylopoulos & Pacherie (2017), appeals to the notion of executable action concepts as well as an intermediate level of action representations, in between P-intentions and M-intentions, called *motor schemas*. Starting with action concepts, we have such concepts that are acquired on the basis of third-personal observation; one might have action concepts of actions they have never performed, simply on the basis of seeing others perform them. Executable action concepts, on the other hand, are action concepts that are formed on the basis of one's own performance of those actions (see, e.g., Quandt & Chaterjee (2015) for a discussion of first-person and third-person action knowledge systems) .For example, someone who has never figure skated before, but is an avid fan of figure skating, may have an observational action concept of a triple salchow, but unless they have the ability to actually perform this action, they do not have an executable action concept of it.

When an executable action concept is deployed in the content of a P-intention, it triggers the appropriate motor representations. On Mylopoulos & Pacherie's proposal, this occurs by way of motor schemas, which form a bridge between P-intentions and M-intentions. Motor schemas are more abstract, stable, and enduring representations of actions than fully specified M-intentions. They store invariant features of action types pertaining to their ordering, spatial and temporal configuration, and relative speed and force. These constitute the general form of an action type, which can be understood as a set of open parameters, each corresponding to an invariant feature of that action. For example, a reaching-towards-target movement might have a corresponding motor schema which includes the following open parameters: hand-to-target distance, target approaching velocity, elbow extension angle, and arm elevation angle (Postachinni et al. 2015). Recent empirical work suggests that these schemas may be stored in the ventro-dorsal stream (the "use" system). When the relevant motor schema is activated in the context of action, this puts constraints on processing by the dorso-dorsal stream, the role of which is to then help "fill in" these parameters yielding fully specified M-intentions where all the details of the action are represented in motoric terms (see discussion in section 3). Here there is also a role reserved for attention in parameterization, in perhaps the way that Wu (2015) envisions when he writes: "What attention is providing to the agent is the relevant target, but this

involves providing content to systems that process that target in a way that sets parameters for and programs an appropriate response" (13). The final step in this account is that motor schemas are both acquired and associated with corresponding executable action concepts by way of a process of bottom-up learning.

Mylopoulos & Pacherie's (2017) account has the advantage of detailing a content-preserving causal process between P-intentions and M-intentions that does not appeal to or presuppose any translation process. But some have complained that too many details are left unspecified. Thus, Shepherd (2017) proposes a solution that has some P-intentions taking *both* propositional and motoric formats: "The solution to the interface problem is that intentions lead a double life. Intentions can take propositionally formatted contents that enable their integration with propositional thought. And intentions have motorically formatted contents that communicate in a fairly direct way with the operations of motoric-level action implementation" (10).

Shepherd (2017) takes the positing of dual-contents for (some) P-intentions to be the best way to explain certain empirical results from sensorimotor adaptation paradigms (e.g., Taylor & Ivry 2011; McDougle et al. 2016). In these experiments, participants are given the task of reaching towards a target. They cannot see their arm, but they are provided with visual feedback on a screen of its trajectory. The experimenters introduce perturbations to this feedback such that participants' movements must be adjusted if they are to successfully reach the target. A well-known finding from these experiments is that the motor system displays drift in the direction opposite the non-veridical sensory feedback, indicating implicit learning that seems aimed at minimizing sensory prediction error. When participants are given the opportunity to compensate for the perturbations by forming explicit strategies to aim away from the proper target and at a neighbouring target instead their reaching errors are completely eliminated at first, but after some time the motor system starts displaying drift once again, this time relative to the aimed at location rather than the proper target. Shepherd takes this type of result to provide evidence for a double-format view of intention, since the intention to aim for the neighbouring target "sets the location around which implicit learning occurs" (8).

Shepherd's proposal is attractive in that it provides a way of smoothing over the gap between propositionally-formatted content and motorically-formatted content by having one and the same state possess both. Still, one might worry that the empirical data that Shepherd finds suggestive indicates only that there is a tight interplay between intentions and the motor system, so that it alone does not tell in favour of a dual content view, as against a view on which intentions and motor representations have different formats, but are nonetheless closely coordinated. In order for the argument to go through, it must be the case that positing dual content is the *best* explanation of the data. But here it is worth noting that Shepherd allows an appeal to deferral and learning can explain the non-accidental link between intentions and the motor system (p.10). He writes:

If an action concept such as GRASPING MY COFFEE MUG becomes linked with the relevant motor schemata via repeated tokening within very similar intentions, there may come a time when all I need to grasp my coffee mug successfully is the tokening of an abstract, propositionally represented intention (e.g., GRASP MY MUG!), which then defers successfully to the associated motor schemata. In such a case, the learning process that links the action

concept with the motor schemata will explain the non-accidental nature of the link.

If this is right, then one might wonder why it is not enough to account for the relevant empirical data without taking on board dual-content intentions. In any case, we do not propose to settle these issues here, we merely wish to highlight open questions that remain for future work.

Conclusion

The DPM model aimed to explicate action production and control in terms of a dynamic hierarchy of intentions. While work in this area is still in progress, a picture on which agentive control depends on rich and flexible interactions among different levels of the hierarchy appears to be well-supported. In this review, we have attempted to clarify and make more precise how the interactions among intentions at various levels of the hierarchy may operate in the light of recent work in philosophy and the cognitive sciences.

Notes

[1.] Another motivation, which we do not have the space to go into here, was explaining where in the mechanisms underlying action control we might locate the states or processes that give rise to the phenomenology of agency, i.e., the distinctive experience or set of experiences that accompany our actions. We will not go into this further here, though it is worth noting that in recent years numerous accounts of agentive phenomenology have been developed in the literature, many of which have been influenced by the DPM model (Bayne 2011; Pacherie 2015; Shepherd 2017; Mylopoulos 2017).

[2.] In addition, it is unclear how motor intentions, as characterized by Brozzo (2017) differ from what Pacherie characterizes as P-Intentions.

[3.] It is also important to note that although vision may be the dominant sensory modality in humans, information from other modalities – e.g., proprioception, touch, audition, or olfaction – is also relevant to action. For instance, grip selection depends on both visual and tactile information (Grafton 2010). Dijkerman and de Haan (2007) have argued for a distinction between perception and action similar to Milner and Goodale's within the somatosensory system, suggested that it corresponds to the distinction between the body image, providing information for body recognition, and the body schema, used in guiding actions towards the self (see de Vignemont 2018 for a discussion of the distinction between body image and body schema). There is little evidence, however, for a similar distinction in other sense modalities.

[4.] To further complicate the picture, Kravitz and colleagues (Kravitz et al. 2011) have identified three pathways respectively supporting spatial working memory, visually guided action and navigation, and Batelli and colleagues (Batteli et al., 2007; 2008) have argued for a fourth 'when' dorsal pathway. This proliferation of dorsal pathways has even led some to question the usefulness of stable pathway models of the organization of visual processing and to argue instead for patchwork or dynamic network models of visual processing (de Haan & Cowey 2011; Galletti & Fattori 2018).

[5.] Recently, Brozzo (2017) has argued that, while the motor representations underlying automatic adjustments in online action guidance are likely inaccessible to consciousness, "there are descriptions of what the agent is doing in more general terms, however, that are represented by motor representations and are also such that the subject may well consciously access that level of description of the action—for instance, *reaching towards the visual target*" (p.21). It is difficult to see, however, how to determine at this point whether it is a motor representation with this content that the participant is aware of, a P-intention with this content, or even a kinaesthetic experience with this content.

[6.] It should be noted that it is unclear which extant views in the literature can actually be characterized as determination views. For example, while Burnston (2017) treats the view that Mylopoulos & Pacherie (2017) defend as a determination view, they maintain that proximal intentions activate corresponding motor schemas, which themselves contain open parameters that must be specified with the help of attention and context, which would seem to favour treating theirs as a biasing view instead.

References

Adam, J. J., Bovend'Eerdt, T. J. H., Schuhmann, T., & Sack, A. T. (2016). Allocentric coding in ventral and dorsal routes during real-time reaching: Evidence from imaging-guided multi-site brain stimulation. *Behavioural brain research*, *300*, 143-149.

Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The 'when'pathway of the right parietal lobe. *Trends in cognitive sciences*, *11*(5), 204-210.

Battelli, L., Walsh, V., Pascual-Leone, A., & Cavanagh, P. (2008). The 'when'parietal pathway explored by lesion studies. *Current opinion in neurobiology*, *18*(2), 120-126.

Bayne, T.   The sense of agency. In F. Macpherson (Ed.), *The Senses* (pp.490-525). Oxford: Oxford University Press.

Binkofski, F., & Buxbaum, L. J. (2013). Two action systems in the human brain. *Brain and Language*, *127*(2), 222-229.

Brand, M. (1984). *Intending and acting*. Cambridge, MA.: MIT Press.

Bratman, M. (1987). *Intention, plans, and practical reason*. Cambridge, MA: Cambridge University Press.

Briscoe, R., & Schwenkler, J. (2015). Conscious vision in action. *Cognitive Science*, 39(7), 1435–1467.

Brogaard, B. (2011). Conscious vision for action versus unconscious vision for action? *Cognitive Science*, 35, 1076–1104.

Brozzo, C. (2017). Motor intentions: How intentions and motor representations come together. *Mind & Language*, *32*(2), 231-256.

Bruno, N., & Franz, V. H. (2009). When is grasping affected by the Müller-Lyer illusion?: A quantitative review. *Neuropsychologia*, *47*(6), 1421-1433.

Bruno, N., Bernardis, P., & Gentilucci, M. (2008). Visually guided pointing, the Müller-Lyer illusion, and the functional interpretation of the dorsal-ventral split: conclusions from 33 independent studies. *Neuroscience & Biobehavioral Reviews*, *32*(3), 423-437.

Burnston, D. (2017). Interface problems in the explanation of action. *Philosophical Explorations, 20*(2), 242-258.

Butterfill, S. & Sinigaglia, C. (2014). Intention and motor representation in purposive action. *Philosophy and Phenomenological Research. 88*(1), 119-145.

Buxbaum, L. J., Kyle, K., Grossman, M., & Coslett, B. (2007). Left inferior parietal representations for skilled hand-object interactions: evidence from stroke and corticobasal degeneration. *Cortex*, *43*(3), 411-423.

Caiani, S. Z., & Ferretti, G. (2017). Semantic and pragmatic integration in vision for action. *Consciousness and cognition*, *48*, 40-54.

Chen, Z. & Saunders, J. A.. (2016). Automatic adjustments toward unseen visual targets during grasping movements. *Experimental Brain Research*, 234(7), 2091-103.

Christensen, W., Sutton, J., & McIlwain, D. J. (2016). Cognition in skilled action: Meshed control and the varieties of skill experience. *Mind & Language*, *31*(1), 37-66.

Cloutman, L. L. (2013). Interaction between dorsal and ventral processing streams: where, when and how?. *Brain and Language*, *127*(2), 251-263.

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy*, *60*(23), 685-700.

Davis, L. H. 1994. Action. in S. Guttenplan (ed.), *A companion to the Philosophy of Mind*, Oxford: Blackwell, pp. 111-117.

de Haan, E. H., & Cowey, A. (2011). On the usefulness of 'what'and 'where'pathways in vision. *Trends in cognitive sciences*, *15*(10), 460-466.

Dijkerman, H. C., & de Haan, E. H. (2007). Somatosensory processing subserving perception and action: Dissociations, interactions, and integration. *Behaviour and Brain Sciences*, 30, 224–230.

Ferretti, G. (2016). Through the forest of motor representations. *Consciousness and cognition*, *43*, 177-196.

Filimon, F., Nelson, J. D., Hagler, D. J. and Sereno, M. I. (2007). Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *Neuroimage*, 37(4), 1315-1328.

Franz, V. H., & Gegenfurtner, K. R. (2008). Grasping visual illusions: Consistent data and no dissociation. *Cognitive Neuropsychology*, *25*(7-8), 920-950.

Fridland, E. (2014). They've lost control: Reflections on skill. *Synthese*, *91*(12), 2729-2750.

Fridland, E. (2017). Skill and motor control: intelligence all the way down. *Philosophical Studies*, *174*(6), 1539-1560.

Galletti, C., & Fattori, P. (2018). The dorsal visual stream revisited: Stable circuits or dynamic pathways?. *Cortex*, 98: 203-217.

Gallese, V. (2007). The "Conscious" dorsal stream: Embodied simulation and its role in space and action conscious awareness. *Psyche*, 13(1), 1–20.

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American psychologist*, *54*(7), 493.

Gollwitzer, P. M., & Sheeran, P. (2006). Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in experimental social psychology*, *38*, 69-119.

Grafton, S. T. (2010). The cognitive neuroscience of prehension: recent developments. *Experimental brain research*, *204*(4), 475-491.

Grafton, S. T., & Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human movement science*, *26*(4), 590-616.

Grünbaum, T. (2017). The Perception-Action Model: Counting Computational Mechanisms. *Mind & Language*, *32*(4), 416-445.

Hamilton, A., & 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*. Oxford: Oxford University Press.

Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioural and Brain Sciences, 24*(5), 849-878.

Jacob, P., & Jeannerod, M. (2003). *Ways of seeing. The scope and limits of visual cognition*. Oxford: Oxford University Press.

Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford: Blackwell.

Jeannerod, M. (2006). *Motor cognition*. Oxford: Oxford University Press.

Jeannerod, M., Decety, J., & Michel, F. (1994). Impairment of grasping movements following a bilateral posterior parietal lesion. *Neuropsychologia, 32*(4), 369-380.

Kalckert, A. & Henrik Ehrsson, H. (2014). The moving rubber hand illusion revisited: comparing movements and visuotactile stimulation to induce illusory ownership. *Consciousness and Cognition, 26*, 117–32.

Kopiske, K. K., Bruno, N., Hesse, C., Schenk, T., & Franz, V. H. (2016). The functional subdivision of the visual brain: Is there a real illusion effect on action? A multi-lab replication study. *cortex, 79*, 130-152.

Kozuch, B. (2015). Dislocation, not dissociation: The neuroanatomical argument against visual experience driving motor action. Mind & Language, 30(5), 572–602.

Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience, 12*(4), 217.

Levy, N. (2017). Embodied savoir-faire: knowledge-how requires motor representations. *Synthese, 194*(2), 511-530.

Mahon, B. Z., & Wu, W. (2015). Cognitive Penetration of the Dorsal Visual Stream. In J. Zeimbekis & A. Raftopoulos (Eds), *The cognitive penetrability of perception: New philosophical perspectives* (pp. 200-217). Oxford: Oxford University Press.

Mahon, B. Z., Kumar, N., & Almeida, J. (2013). Spatial frequency tuning reveals interactions between the dorsal and ventral visual systems. *Journal of cognitive neuroscience, 25*(6), 862-871.

McDougle, S.D., Ivry, R.B., and Taylor, J.A. (2016). Taking aim at the cognitive side of learning in sensorimotor adaptation tasks. *Trends in Cognitive Sciences, 20*(7), 535-544.

McIntosh, R. D., & Lashley, G. (2008). Matching boxes: Familiar size influences action programming. *Neuropsychologia, 46*(9), 2441-2444.

Mele, A. R. (1992). *Springs of action: Understanding intentional behavior*. Oxford: Oxford University Press.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia, 46*, 774–785.

Milner, A., & Goodale, M. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Mole, C. (2009). Illusions, demonstratives, and the zombie action hypothesis. *Mind, 118*(472), 995-1011.

Mylopoulos, M. & Pacherie, E. (2017). Intentions and Motor Representations: The Interface Challenge. *Review of Philosophy and Psychology, 8*(2), 317-336.

Mylopoulos, M. (2017). A cognitive account of agentive awareness. *Mind & Language, 32*, 545-563.

Nanay, B. (2013a). *Between perception and action*. Oxford University Press.

Nanay, B. (2013b). Is action-guiding vision cognitively impenetrable? In Proceedings of the 35th annual conference of the cognitive science society (pp. 1055–1060). Hillsdale, NJ: Lawrence Erlbaum.

Niessen, E., Fink, G. R., & Weiss, P. H. (2014). Apraxia, pantomime and the parietal cortex. *NeuroImage: Clinical, 5*, 42-52.

Pacherie, E. (2000). The content of intentions. *Mind & Language, 15*(4), 400–432.

Pacherie, E. (2006). Towards a dynamic theory of intentions. In S. Pockett, W. P. Banks, & S. Gallagher (Eds.), *Does consciousness cause behavior? An investigation of the nature of volition* (pp. 145–167). Cambridge, MA: MIT Press.

Pacherie, E. (2007). The sense of control and the sense of agency. *Psyche*, 13(1), 1–30.

Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107(1), 179–217.

Pacherie, E. (2011). Non-conceptual representations for action and the limits of intentional control. Social Psychology, 42, 67–73.

Pacherie, E. & Haggard, P. (2010). What are intentions?. In, L. Nadel & W. Sinnott-Armstrong (eds.) *Conscious Will and Responsibility. A tribute to Benjamin Libet* (pp. 70-84). Oxford: Oxford University Press.

Pacherie, E. (2015). Time to act: The dynamics of agentive experiences. In B. Eitam & P. Haggard (Eds.), The sense of agency. Oxford: Oxford University Press.

Papineau, D. (2015). Choking and the yips. *Phenomenology and the Cognitive Sciences*, 14, 295-308.

Pisella, L., Binkofski, F., Lasek, K., Toni, I., & Rossetti, Y. (2006). No double-dissociation between optic ataxia and visual agnosia: Multiple sub-streams for multiple visuo-manual integrations. *Neuropsychologia*, 44(13), 2734-2748.

Postacchini, R., Paoloni, M., Carbone, S., Fini, M., Santilli, V., Postacchini, F., Mangone, M. (2015). Kinematic analysis of reaching movements of the upper limb after total or reverse shoulder arthroplasty. *Journal of Biomechanics, 48*(12), 3192-3198.

Prestwich, A., Sheeran, P., Webb, T. L., & Gollwitzer, P. M. (2015). Implementation intentions. *Predicting health behavior*, 321-357.

Quandt, L. C., & Chatterjee, A. (2015). Rethinking actions: implementation and association. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(6), 483-490.

Rizzolatti, G., & Matelli, M. (2003). Two different streams form the dorsal visual system: Anatomy and functions. *Experimental Brain Research*, 153, 146–157.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3(2), 131-141.

Schenk, T., & McIntosh, R. D. (2010). Do we have independent visual streams for perception and action? *Cognitive Neuroscience*, 1, 52–78.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge university press.

Shepherd, J. (2014). The contours of control. *Philosophical Studies*, 170(3), 395–411.

Shepherd, J. (2015). Conscious control over action. *Mind & Language, 30*(30), 320-344.

Shepherd, J. (2016). Conscious action/zombie action. *Noûs, 50*(2), 419-444.

Shepherd, J. (2017). Skilled action and the double life of intention. Philosophy and Phenomenological Research, doi:10.1111/phpr.12433.

Shepherd, J. (2017). The experience of acting and the structure of consciousness. *The Journal of Philosophy*, 114(8), 422-448.

Taylor, J. & Ivry, R.B. (2011). Flexible cognitive strategies during motor learning. *PLoS Computational Biology 7*(3), pe1001096.

van Elk, M., van Schie, H., & Bekkering, H. (2014). Action semantics: a unifying conceptual framework for the selective use of multimodal and modality-specific object knowledge. *Physics of life reviews*, 11(2), 220-250.

De Vignemont, F. (2018). *Mind the Body*. Oxford: Oxford University Press.

Vry, M. S., Tritschler, L. C., Hamzei, F., Rijntjes, M., Kaller, C. P., Hoeren, M., Umarova, R., Glauche, V., Hermsdoerfer, J., Goldenberg, G., Hennig, J., & Weiller, C. (2015). The ventral fiber pathway for pantomime of object use. *NeuroImage, 106*, 252-263.

Wittgenstein, L. (1953). *Philosophical Investigations*. Oxford: Blackwell.

Wong, H. Y. (2015). On the significance of bodily awareness for bodily action. *The Philosophical Quarterly*, 65(261), 790-812.

Wu, W. (2013). The case for zombie agency. *Mind, 122*(485), 217-230.

Wu, W. (2014). *Attention*. New York, NY: Routledge.

Wu, W. (2014). Against division: Consciousness, information and the visual streams. *Mind & Language*, 29(4), 383–406.

Wu, W. (2016). Experts and deviants: the story of agentive control. *Philosophy and Phenomenological Research*, *93*(1), 101-126.