Motor intentionality
Elisabeth Pacherie

To cite this version:
Motor intentionality

Elisabeth Pacherie

Institut Jean Nicod, CNRS UMR 8129, Institut d’Etude de la Cognition, École Normale Supérieure & PSL Research University, Paris, France.

Abstract:

In his book, *Phenomenology of Perception*, the French philosopher Maurice Merleau-Ponty first coined the phrase 'motor intentionality'. At the same time he highlighted the contrast between motor and cognitive intentionality, he also emphasized their generally smooth interplay in normal agents. An account of motor intentionality should thus aim at elucidating not just what distinguishes motor intentionality from more cognitive forms of intentionality but also how motor intentionality relates to these more cognitive forms of intentionality. Using Merleau-Ponty's discussion of motor intentionality as my starting point, I then consider how more recent conceptual and empirical work can help sharpen our understanding of the distinctiveness of motor intentionality. In contrast to Merleau-Ponty, I defend a representational stance on motor intentionality. Finally, I turn to the challenges raised by its interplay with more cognitive forms of intentionality and the problem of explaining how our motor behavior can be responsive to our intentions.

Keywords: motor intentionality; motor representations; intentions; non-conceptual content; visual pathways; apraxia; interface problem; motor schemas.
1. Introduction

In his famous book *Phenomenology of Perception*, first published in 1945, the French philosopher Maurice Merleau-Ponty coined the phrase 'motor intentionality' using it to refer to the form of intentionality exemplified by purposive, skillful, unreflective bodily activities, as opposed to the more cognitive, conceptual and representational forms of intentionality typical of conscious intentions. He introduced this notion in his long discussion of the case of Schneider, a soldier in the German Army who suffered serious brain injuries during World War I and displayed a large number of neuropsychological impairments. Merleau-Ponty used Schneider's case to highlight the contrast between motor and cognitive intentionality but also to emphasize their generally smooth interplay in normal agents. In what follows, I will explore this contrast and interplay. How should we characterize motor intentionality? Is it best described, as Merleau-Ponty would have it, as a form of non-representational intentionality? If not, how do motor representations differ from the representations involved in conscious intentions? How can motor intentionality and more cognitive forms of intentionality be integrated?

In section 2, I take Merleau-Ponty's discussion of Schneider's case as my starting point. In section 3, I consider more recent conceptual and empirical work that can help elucidate the distinction between motor and cognitive intentionality but also shed light on the challenges raised by their interplay. In section 4, I defend a representational stance of motor intentionality and discuss the format and contents of motor representation. Finally, section 5 will discuss the interplay between motor and cognitive intentionality and the problem of explaining how our motor behavior can be responsive to our intentions?

2. Merleau-Ponty on motor intentionality
In *Phenomenology of Perception*, Merleau-Ponty used the case of Schneider to motivate the need to posit motor intentionality as a basic form of intentionality. Schneider, a soldier in the German army in World War I, suffered serious brain injuries when wounded by the explosion of a mine. He became a patient of the psychologist Adhémar Gelb and the neurologist Kurt Goldstein, who in their case reports described the large array of neuropsychological impairments he displayed, including alexia, form agnosia, loss of movement vision, loss of visual imagery, tactile agnosia, loss of body schema, loss of position sense, acalculia and loss of abstract reasoning (Goldstein & Gelb 1918; Goldstein 1923). Merleau-Ponty was especially interested in Schneider's pattern of performance in different motor tasks, as described by Gelb and Goldstein. Schneider presented a dissociation between a preserved ability to perform what Gelb and Goldstein termed 'concrete movements' and an impaired ability to perform 'abstract movements'. In their terminology, concrete movements correspond to habitual movements performed in everyday life and abstract movements are isolated, arbitrary movements not relevant to any actual situation, such as moving arms and legs to order, or bending and straightening a finger. For instance, Schneider could grasp his nose with

---

1 The validity of Schneider’s case and the exact nature of his impairments have been a matter of debate. Goldenberg (2003) argues that Gelb and Goldstein’s minds "were clouded by the enthusiasm of proving the truth of an all-embracing theory of the human mind and its reaction to brain damage" (2003: 292), leading them to embellish their description of the case, while comforted in their enthusiasm by a patient eager to please. Others, however, have pointed out that aspects of Schneider’s behavior that raised Goldenberg’s suspicions, such as the compensation of visual form agnosia by kinaesthetic mediation, are modes of compensation spontaneously used by patients with similar deficits (Farah 2004; Marotta and Behrmann 2004). It is also a matter of debate whether Schneider’s case should be classified as an example of apperceptive visual agnosia or rather of integrative agnosia (Marotta and Behrmann, 2004). Importantly for present purposes, even if doubts are likely to persist regarding the validity of all of Goldstein and Gelb’ claims about their patient, the dissociation between identification and localization, that is the focus of Merleau-Ponty’s discussion is now well documented, as will be discussed in section 3, and has been found not just for the visual modality but also for the tactile modality (Paillard et al. 1983).
his hand but not point to it; nor could he interrupt his grasping movement midway on order or touch his nose with a ruler. He could perform habitual actions with speed and precision, like taking a match out of a box and lighting a lamp, but was at a loss when asked to perform an abstract, arbitrary movement, like drawing a circle in the air with his arm. Finally, he could perform or pantomime habitual movements on order, but only by placing himself mentally in the actual situation to which they correspond and then executing them in perfect detail.

Taking the dissociation between Schneider's inability to point to his nose and his preserved ability to grasp his nose as evidence in support of a distinction between cognitive and motor intentionality, Merleau-Ponty wrote:

It must therefore be concluded that 'grasping' or 'touching', even for the body, is different from 'pointing'. From the outset the grasping movement is magically at its completion; it can begin only by anticipating its end, since to disallow taking hold is sufficient to inhibit the action. And it has to be admitted that a point on my body can be present to me as one to be taken hold of without being given in this anticipated grasp as a point to be indicated. But how is this possible? If I know where my nose is when it is a question of holding it, how can I not know where it is when it is a matter of pointing to it? It is probably because knowledge of where something is can be understood in a number of ways. (Merleau-Ponty 2002, p. 119)

Merleau-Ponty proposed that this dissociation points to the existence of different ways of knowing or understanding locations in space. Pointing to one's nose demands that one be able to form a representation of the positions of one's nose and hand in objective space. In contrast, grasping one's nose involves a practical understanding of bodily space, "where the patient is conscious of his bodily space as the matrix of his
habitual action, but not as an objective setting” (Merleau-Ponty 2002, p. 119). Merleau-Ponty also emphasized the independence of this practical understanding from an objective understanding of bodily space. He wrote:

A patient of the kind discussed above, when stung by a mosquito, does not need to look for the place where he has been stung. He finds it straight away, because for him there is no question of locating it in relation to axes of co-ordinates in objective space, but of reaching with his phenomenal hand a certain painful spot on his phenomenal body, and because between the hand as a scratching potentiality and the place stung as a spot to be scratched a directly experienced relationship is presented in the natural system of one’s own body. (Merleau-Ponty 2002, p. 121).

Importantly, this practical understanding is not confined to one's bodily space narrowly conceived and to actions directed at one's body. This system also encompasses the surrounding space and the familiar objects it contains, offering themselves as poles of action in relation to the body's potentialities. Thus, according to Merleau-Ponty: "In the action of the hand which is raised towards an object is contained a reference to the object, not as an object represented, but as that highly specific thing towards which we project ourselves, near which we are, in anticipation, and which we haunt " (Merleau-Ponty 2002, p. 159)

Here, Merleau-Ponty appears to take the dissociation between different types of motor tasks in Schneider's case as evidence for the existence of a way of being directed towards one's body and towards objects in one's surroundings that functions independently of conceptual representations of their locations in objective space. He seems to claim both that motor intentionality is preserved in pure form in Schneider and
also, more generally, that motor intentionality is our normal way of relating to our body and surroundings and what enables our unreflective, skillful goal-directed activities.

There is, however, another line of argumentation that runs simultaneously in Merleau-Ponty's long discussion of Schneider's case and leads to a conflicting conclusion. Schneider is unable to draw a circle in the air in the normal way: "asked to trace a square or a circle in the air, he first 'finds' his arm, then lifts it in front of him as a normal subject would do to find a wall in the dark and finally he makes a few rough movements in a straight line or describing various curves, and if one of these happens to be circular he promptly completes the circle" (Merleau-Ponty 2002, p. 126). Since the patient can move, he doesn't lack motility and since he can recognize when the movements he makes happen to be circular, he doesn't lack a representation of the movement. Here, Merleau-Ponty concludes that what he lacks is "something which is an anticipation of, or arrival at, the objective and is ensured by the body itself as a motor power, a 'motor project' (Bewegungsentwurf), a 'motor intentionality' in the absence of which the order remains a dead letter" (Merleau-Ponty 2002, p. 127).

As several authors have pointed out and as Jensen (2009) discusses in detail, these two lines of reasoning suggest there is at best an ambiguity and at worst an inconsistency in Merleau-Ponty's interpretation of Schneider's. On the one hand, he appears to claim that pure motor intentionality is preserved in Schneider's case, but, on the other hand, he also appears to take his inability to convert the thought of a movement into actual movement as evidence of an impairment of motor intentionality. How can motor intentionality be claimed both to be preserved and to be impaired in the same person? Unless they are qualified, the two claims are clearly inconsistent. However, they might be reconciled if we consider that for Merleau-Ponty, motor intentionality is both (i) a basic form of intentionality, distinct from, and capable of
functioning independently of, more abstract, conceptual, objective representational forms of intentionality and (ii) a form of intentionality that also insures the transition between more abstract forms of intentionality (e.g., thoughts about movement) and actual movements. Merleau-Ponty (2002, pp. 127-128) contrasts concrete movement as centripetal and having as background the world as given and abstract movement as centrifugal and as constructing its own background and projecting it, or throwing it out, on the world. Importantly, he takes motor intentionality to be what makes possible both abstract and concrete movements. In a way, then, motor intentionality itself has both a 'centripetal’ dimension, where, as stated in claim (i), it can operate independently of more abstract forms of intentionality, and a 'centrifugal’ dimension where it serves a function of projection of abstract movements into the world, in accordance with claim (ii). Thus, if we understand Merleau-Ponty as suggesting that, in Schneider's case, the centripetal dimension of motor intentionality is preserved, while its centrifugal dimension is impaired, the threat of inconsistency might be avoided. This also means that an account of motor intentionality should aim at elucidating not just what distinguishes motor intentionality from more cognitive forms of intentionality but also how motor intentionality relates to these more cognitive forms of intentionality. In particular, such an account should try and spell out what exactly the function of projection ascribed to motor intentionality by Merleau-Ponty involves and thus move beyond his own largely metaphorical description of this function.

3. Motor intentionality as a basic form of intentionality: empirical evidence

According to Merleau-Ponty, motor intentionality constitutes a basic form of intentionality, distinct from more cognitive forms of intentionality and capable of functioning independently of them. Findings from several lines of empirical research in
cognitive science and neuroscience appear to support the distinction and dissociability of motor intentionality and other forms of intentionality. In particular, a large body of empirical evidence ranging from electrophysiological studies of macaque monkey brains to neuropsychological studies of patients with brain damage and behavioral studies in healthy humans support a dual model of visual processing, with a visuomotor system subserving the visual guidance of actions directed at objects in the environment (vision-for-action) and a visual perceptual system subserving the construction of visual percepts and conscious object perception (vision-for-perception).

In the early 1980’s, neuro-anatomists and physiologists established the existence of two separate cortical pathways, ventral and dorsal, subserving different functions in the visual cortex of primates. (Ungerleider & Mishkin 1982). In-depth studies of patients with lesions in either the dorsal or the ventral pathways provided evidence that processing in the ventral pathway supports 'vision-for-perception' while processing in the dorsal pathway supports 'vision-for-action'. The most famous and widely discussed evidence is probably Milner and Goodale's analysis of patient D.F. (Milner & Goodale 1995). As a consequence of carbon monoxide poisoning, D.F. suffered important lesions of the ventral pathway. As a result, she had visual form agnosia. D.F. is described by Milner and Goodale as unable to recognize everyday objects, to visually identify simple shapes or to tell whether two visual shapes are the same or different. Yet her visuomotor abilities appeared intact. She could reach out and pick up objects with remarkable accuracy, shaping her hand optimally for the grip. When asked to post a card through a slit, she oriented the card correctly, despite being at chance when asked to report the orientation of the slit. In contrast to D.F., patient A. T., studied by Jeannerod

---

2 See Jacob and Jeannerod (2003) for a detailed discussion of this evidence and assessment of its significance.
and colleagues (Jeannerod et al. 1994), had a lesion of the dorsal stream and suffered from optic ataxia. A.T.'s perception of the shape, size and orientation of objects were normal, but her grasping movements directed at objects were systematically incorrect. The co-existence in D.F. of impaired conscious visual perception and object recognition and of preserved visuomotor abilities and the inverse dissociation found in A.T. suggest that visuomotor representations need not be derived from conscious visual perceptions but can be built independently. These dissociations also suggest that conscious visual representations cannot be directly derived from intact visuomotor representations.

Finally, psychophysical experiments in healthy human adults have also shown a dissociation between the processing responsible for accurate visuomotor processing for pointing or grasping and the processing responsible for perceptual awareness. For instance, Bridgeman and colleagues (Bridgeman et al. 1979) conducted a series of experiments that exploited the phenomenon of saccadic suppression. During saccades, i.e., rapid eye movements, vision is partially suppressed and changes in the positions of objects in the visual field are not consciously perceived. Bridgeman and colleagues instructed the participants to point to a target that had just been displaced and extinguished. On some of the trials, the displacement occurred during saccades, preventing the participants to perceive the target displacement. Bridgeman and colleagues found that the accuracy of pointing was not affected by conscious detection or failure to detect the target displacement. In a later set of experiments, Bridgeman and colleagues (Bridgeman et al. 1981) used the dot in frame illusion, where a stationary dot set against a large undifferentiated background moving in one direction appears to be moving in the opposite direction. They found again that although their perceptual judgments of the position of the dot were affected by the dot's apparent motion, their pointing accuracy wasn't. These experiments suggest that visual awareness of the
position and motion of a target and visually guided pointing at a target are largely independent processes. Similarly, size-contrast illusions have been shown to affect conscious perception and judgment but not grasping performance. The Titchener illusion (also known as the Ebbinghaus illusion) is a display consisting of two circles of equal size, one surrounded by annulus of smaller circles, the other surrounded by larger circles. As a result, the circle surrounded by smaller circles is perceived as larger than the other central circle. Aglioti et al. (1995) used a three-dimensional version of the illusion using plastic disks and had their participants make a perceptual judgment and pick up one of the two central disks. A grasping movement involves a progressive opening of the grip where the fingers stretch up to a maximum aperture, followed by a closure of the grip until it matches object size. Maximum grip aperture occurs at about 60% to 70% of the duration of the movement and is reliably correlated with the object's size (Jeannerod 1981). Aglioti and colleagues used this property of the motor grasping pattern as an index of the computation of the object size made by the visuomotor system. They found that while perceptual judgments about object size were affected by the illusion, the grip wasn't and remained correlated with the object's actual size.3

Similar findings regarding pointing and grasping have been reported for a variety of other visual illusions including the Müller-Lyer illusion (Daprati & Gentilucci 1997),

---

3 The design of these experiments has raised certain methodological criticisms. For instance, Franz et al. (2000) argued that their results might be due to an asymmetry between the perceptual and the motor task (the perceptual task requiring the subjects to compare two discs, whereas in the motor task they could focus their attention on a single disc), and as such provided no evidence for a dissociation between perception and action. However, Haffenden and Goodale (1998) obtained similar results in a modified version of the task where this asymmetry was not present and the motor task and perceptual task were matched. For discussions of the methodological issues concerning illusion studies and of the degree to which they support the dual visual system hypothesis, see Jacob & Jeannerod (2003) and Briscoe (2008, 2014).
the Ponzo illusion (Jackson & Shaw 2000), the Kanizsa compression illusion (Bruno and Bernardis 2002) and the Hollow-Face illusion (Króliczak et al. 2006). In each case, there is a divergence between what subjects consciously see and their visually guided behavior, suggesting that the spatial information used for visually-guided action and the (illusory) spatial content of conscious visual experience might be processed relatively independently.

It is important to note, however, that our understanding of the visual pathways has evolved considerably since Milner and Goodale (1995) proposed their dual-system model. Substantial evidence has accrued that the anatomical and functional separation between the dorsal and ventral pathways is far from complete, casting doubt of the validity of a simple dissociation between vision for perception and vision for action and suggesting instead a more complex organization of the visual processing. Thus, Rizzolatti and Matelli (2003) have described two anatomically segregated sub-circuits of the dorsal stream, a dorso-dorsal pathway and a ventro-dorsal pathway. It has been proposed that the dorso-dorsal pathway is concerned with immediate visuo-motor control and the ventro-dorsal pathway with the long-term storage of the particular skilled actions associated with familiar objects, with lesions to one or the other pathways leading to different neuropsychological impairments (Binkofski & Buxbaum 2013; Pisella et al. 2006).

In addition, neuroanatomical studies have uncovered many connections between the dorsal substreams and the ventral ventral stream, indicating that these streams are able to communicate with each other in a bidirectional way and suggesting that the ventro-dorsal substream may constitute an interface between the ventral and the dorsal streams of visual information processing. Similarly brain imaging studies indicate that the dorsal and ventral streams are often jointly involved in grasping, notably in
situations involving delayed or pantomimed grasping, situations when information about the object from pictorial cues or memory is needed to control the grasping movement, or tool use where conceptual knowledge needs to be to allow for the selection of the most appropriate grasp (for reviews, see Cloutman 2013; Grafton 2010). Thus, on the one hand, evidence of dissociations between visuo-motor processing and visual perception processing appear to support Merleau-Ponty’s contention that motor intentionality constitutes a basic form of intentionality, distinct from more cognitive forms of intentionality and capable of functioning independently of them. On the other hand, evidence of substantial crosstalk between streams appears consistent with his further contention that motor intentionality insures the transition between more abstract forms of intentionality (e.g., thoughts about movement) and actual movements.

Before I consider the challenges raised by the interfacing of motor intentionality and more cognitive forms of intentionality, let me try to offer first a fuller characterization of motor intentionality.

4. Motor representations

Merleau-Ponty characterizes motor intentionality as non-representational, whereas cognitive scientists are generally happy to talk of the dorsal pathway as computing sensorimotor representations. Is this just a matter of terminological sloppiness on the part of cognitive scientists or is instead Merleau-Ponty’s use of the term "representation" highly loaded and perhaps overly restrictive?

For something to qualify as a representation in Merleau-Ponty’s sense it must have propositional, conceptual content, and represent an object or a situation in an objective or detached fashion. However, many cognitive scientists and philosophers currently operate with a less demanding notion of representation. For instance, according to the
account proposed by Bermúdez (1998), for a state to qualify as representational the following criteria should be met: (1) the state should have correctness conditions and allow for the possibility of misrepresentation; (2) it should be compositionally structured; (3) it should admit of cognitive integration; and (4) it should play a role in the explanation of behavior that cannot be accounted for in terms of invariant relations between sensory input and behavioral output. This characterization leaves it open whether a representation has conceptual content or not, whether its content is objective or detached or not and whether its format is propositional. Importantly, both cognitive integration and compositionality are graded notions. So, one way of drawing the distinction between conceptual and non-conceptual representations would be to say that conceptual representations must satisfy more stringent criteria of full cognitive integration and full compositionality. Indeed, Bermúdez suggests that the distinction between conceptual and non-conceptual content may in part be a matter of degree of compositionality and cognitive integration.4

Format and content of motor representations

Several authors (Butterfill & Sinigaglia, 2014; Jacob & Jeannerod, 2003, Pacherie, 2000, 2011) have argued that sensorimotor representations, like perceptual representations, have non-conceptual content, but also that this non-conceptual content is of a different kind from the non-conceptual content of perception. According to these authors, a motor representation represents the goal of an action in a specific non-

4 Bermúdez (1998) also argues, perhaps more contentiously (see Levine 2001) that there is a constitutive link between a capacity for conceptual thought and a capacity for genuine inference, where having a capacity for genuine inference is linked to an ability to appreciate the rational grounds for, and thus to justify, one’s inferences, and that capacity for justification requires language mastery.
conceptual format. This representation of the goal of an action (say, reaching for an object) is not just a representation of the target object toward which the action is directed, it also include a representation of the final state of the acting body when that object has been reached. In simple, object-oriented actions (i.e., when an object is the target of an action), the visual attributes of this object are represented in a specific, 'pragmatic' mode used for the selection of appropriate movements and distinct from other modes of representation used for other aspects of object-oriented behavior (categorization, recognition, etc.). In that sense, pragmatic representations are not as informationally rich as perceptual representations, since they represent objects attributes only to the extent that they are relevant to the selection of motor patterns. Jeannerod (1997) suggests that the function of these representations ‘falls between’ a sensory function (extracting from the environment attributes of objects or situations relevant to a given action) and a motor one (encoding certain aspects of that action). In other words, these representations should be viewed as relational, with the body and the target object functioning as the terms of the relation. What they represent are neither states of the body per se nor states of the environment per se, but rather relations between body and goal. To use a different formulation, we could say that the goal is given under a specific mode of presentation; it is represented in terms of the motor patterns that it affords to the agent.

Another important aspect of motor representations is their dynamical character: they do not just represent relations between body and goal, they represent dynamic relations between them. This characteristic is linked to their role in the guidance and control of the action as it unfolds. In order for a motor representation to guide an action, it must anticipate the future states of the environment and of the acting body itself; in order to control it it must allow for adjustments during execution. In recent decades,
theories of motor control have emphasized the role of internal forward or predictive models. These models capture the causal relationships between motor acts and their sensory consequences and can be used by the motor system to estimate the effects of the motor commands sent to the effectors, compare these predicted effects with sensory feedback, and make adjustments if needed (for full descriptions of these models, see Desmurget & Grafton 2000; Wolpert, Ghahramani & Jordan 1995; Wolpert & Kawato 1998). The content of motor representations is thus dynamical in the sense both that it gets elaborated over time — it becomes more determinate through feedback — and that the motor representation is itself responsible for making available the information that will make the content more determinate. For instance, to adjust one's grip on an object, one needs accurate information about its weight, compliance and surface texture and sensory feedback will be needed to adjust initial estimates, but for sensory feedback to become available one needs to grasp the object in the first place.

**Are motor 'representations' really representations?**

One may agree that motor intentionality operates along the lines just described, but still be skeptical that the concept of representation plays an explanatory role here and contend instead that motor intentionality is better characterized non-representationally in terms of dynamic systems of self-organizing continuous reciprocal causation between sensorimotor processes and the environment (e.g., Dreyfus 2000; Gallagher 2008).

In the remainder of this section, I argue that motor 'representations' meet Bermúdez’ criteria for representationality.

Bermúdez’ first criterion for a state to count as representational is that it have correctness conditions. One important characteristic of motor representations is their
Janus-faced structure, their function falling between a sensory function and a motor one. A motor representation represents a situation as affording a certain goal, and it does so by representing the motoric means by which the goal is to be achieved. For instance, it represents an object as reachable by representing how the reaching is to be effected. As a result, the classical distinction between states with a mind-to-word direction of fit, and states with a world-to-mind direction of fit (e.g., Searle 1983) while useful as a way of contrasting states such as beliefs and desires does not easily apply to motor representations. Rather, motor representations may be seen as akin to what Millikan (1995) calls 'pushmi-pullyu' representations (or PPRs), that is, hybrid representations with a dual direction of fit. PPRs, according to Millikan (1995), are not simply conjunctions of a descriptive plus a directive representation; rather they are more primitive and computationally less demanding than either purely descriptive or purely directive representations. If we accept that motor representations have this hybrid character, this should be reflected in their correction conditions. A motor representation of an object as to be reached by such and such motoric means, would have dual correctness conditions. For it to be correct it would have to be the case both that the object in question is indeed reachable by these motoric means and that it actually be reached by these motoric means. This characterization of the correctness conditions of motor representations also makes sense of the idea that the success of an action does not just depend on the fact that a certain outcome is achieved, but also on the specific way in which the outcome is achieved. For a given motor representation to be correct, it is not sufficient that it causes some series of changes in the relations between body and world, where the last element in the series corresponds to some desired outcome, the changes must also conform to a certain dynamical pattern.
Motor representations also have structure and exhibit some form of compositionality, thus meeting Bermúdez' compositionality criterion. They have identifiable constituent units (e.g., reaching, grasping, rotating, lifting, transporting, releasing) that can be combined in various ways. Different actions will involve different combinations of these and other categories of units and, at a higher level of organization, more complex actions will in turn involve combinations of relatively simple actions such as putting an object in a container. For instance, this action could be a recurrent element in the complex action of packing my suitcase before a trip.

In addition, motor representations do not just have a lexicon; they also have what may be called a 'grammar' for assembling the constituent units into a coherent pattern. There are spatial, temporal, and motor (kinematic and biomechanical) constraints on the coordination of action that must be reflected in this 'grammar'. The coordination of reaching and grasping, some aspects of which were already briefly mentioned in the previous section, may serve as an illustration. First, the combination of reaching and grasping units must obey certain spatial constraints. Reaching is mostly achieved by the proximal joints of the arm and makes use of an egocentric or body-centered system of representation of locations. Grasping on the other hand is a function of the intrinsic shape and size of the target object; it involves a transformation of visual information encoded in allocentric, object-centered coordinates into motor information encoded in the system of coordinates used to define the prehension space. Yet reaching and grasping must be spatially compatible. In particular, reaching must take into account not just the location of the object but also its orientation, so that the final position of the arm is compatible with the correct position of the hand and fingers for grasping the object. Second, reaching and grasping must also be temporally coordinated. As we already mentioned, their temporal coordination goes beyond mere succession. The fingers begin
to shape during transportation of the hand at the object location. Maximum grip aperture occurs at about 60% to 70% of the duration of the movement and is reliably correlated with the object’s size. Third, a motor representation normally codes for transitive movements, where the goal of the action determines the global organization of the motor sequence. For instance, the type of grip chosen for a given object is a function not just of its shape and size but also of the intended activity. For instance, the same object may be held with a precision grip or with a power grip depending on whether I intend to put it in a large box or to insert it in a tight-fitting container. Similarly, the same cup will be seized in different ways depending on whether one wants to carry it to one’s lips or turn it upside down. Finally, the biomechanical constraints and the kinematic rules governing the motor system are also reflected in motor representations. Bodily movements as represented in motor representations respect the isochrony principle (the tangential velocity of movements is scaled to their amplitude), Fitt’s law (the time required to rapidly move to a target area is a function of the ratio between the distance to the target and the width of the target) or the two-third power law between curvature and velocity.

Motor representations also admit of cognitive integration (Bermúdez’ third criterion). As we have just seen, how an object is grasped is a function not just of its size, shape and orientation, but also of what we intend to do with it. In addition, how we interact with an object also depends on its function, where the function may not be visually salient. Thus, motor representations will be influenced by knowledge of function. More generally, our motor interactions with an object will often be determined not only by sensory information immediately available to the agent but also by her stored beliefs and knowledge regarding certain attributes and properties of the object (for instance, I may know from previous experience that this pot is heavier than it
Motor representations also connect up with our motivational states. We do not blindly respond to all the solicitations for action that the environment provides. Which motor representations are formed and acted upon is not just a function of environmental saliencies; it can be determined in part by the agent's motivational states, her higher-order goals and intentions or her emotional states (Pacherie, 2002). Motor representations are thus cognitively penetrable to a certain extent and can be influenced by information coming from other sources.

Moreover, it may be argued that the cognitive integration of motor representations is not just a matter of motor representations being influenced by other cognitive states. The influence can also work in the other direction. In particular, there is evidence that motor representations may be activated not just when we prepare to act but also when we observe others acting. In the last two decades neurological studies have yielded a set of important results on mirroring processes. In a series of single neuron recording experiments on macaque monkeys investigating the functional properties of neurons in area F5, Rizzolatti and his colleagues discovered so-called mirror-neurons, i.e. sensorimotor neurons that fire both during the execution of purposeful, goal-related actions by the monkey and when the monkey observes similar actions performed by another agent (for reviews, see Rizzolatti & Craighero 2004; Rizzolatti & Sinigaglia 2008). In addition, a large body of neuroimaging experiments have investigated the neural networks engaged during action generation and during action observation in humans, revealing the existence of an important overlap in the cerebral areas activated in these two conditions (for reviews, see Grèzes & Decety 2001; Jeannerod 2006). The existence of such a mirror system in humans is also supported by behavioral experiments on motor interference, where observation of a movement is shown to
degrade the performance of a concurrently executed incongruent movement (Brass, Bekkering & Prinz 2001; Kilner, Paulignan, & Blakemore 2003).

These results have been interpreted as support for the existence of a process of motor simulation or motor resonance whereby the observation of an action activates in the observer a motor representation of the action that matches the motor representation activated in the brain of the agent. Once a match is established, it enables the observer to apply predictive models in his or her motor system to interpret observed movements and to infer their goal. Thus, motor representations may contribute at least certain premises to cognitive systems engaged in the interpretation of intentional behavior.5

Finally, the existence of a bidirectional link between the processing of linguistic items pertaining to action concepts and the activation of motor representations is also well documented. Thus, passively reading action verbs has been found to somatotopically activate areas of the motor and premotor cortex associated with the relevant body parts needed to carry out the specified actions (Hauk et al. 2004). For example, the different patterns of activation found in the motor cortex when reading the words 'kick', 'pick' or 'lick' overlap significantly with the actual activation that takes place when carrying out these actions with the relevant effectors of foot, hand, and mouth respectively. Conversely, stimulation of the motor system has been found to affect the linguistic processing of action concepts. For instance, one study found that applying TMS to hand and foot areas of the motor cortex improved the recognition of

---

5 The nature and extent of the understanding of others mirroring processes can provide has given rise to an intense debate, with some theorists seeing them as the fundamental neural basis of human social cognition (e.g., Gallese 2007), while others hold more deflationary views (e.g., Jacob 2008).
hand-related (‘pick’) and foot-related (‘kick’) action verbs respectively in lexical decision tasks (Pulvermüller et al. 2005; see also Kiefer & Pulvermüller 2012).

The last criterion to be considered is explanatory usefulness. For motor representations to be vindicated, it must also be demonstrated that a purely mechanical explanation of the motor behavior would not do. According to Bermúdez, the need for explanations appealing to contentful states arises in situations where the behavior to be explained cannot be accounted for in terms of invariant relations between sensory input and behavioral output. Our discussion of the influence of cognitive and motivational factors on the construction of motor representations should make it clear that the motor behavior they are meant to explain could not be explained in terms of a lawful correlation between sensory stimulus and behavioral response. For instance, the same sensory stimulus (a horizontal bar in front of the agent) will be responded to with either an overhand or an underhand grip depending on what the agent intends to do. A mechanistic explanation may perhaps be enough to account for reflexes, but the movements we want to explain are relationally characterized movements — movements related to a certain goal - and as Bermúdez (1998, p. 86) suggests, for such movements we need intentional explanations.

5. The interplay of motor and cognitive intentionality.

As we saw in section 2, there are two lines of arguments in Merleau-Ponty’s discussion of Schneider’s case. In what Jensen (2009) calls the argument from concrete behavior, Merleau-Ponty appears to claim that motor intentionality is preserved in Schneider’s case and is what enables him to perform concrete movements. In contrast, in the argument from abstract behavior, he seems to claim that motor intentionality is impaired in Schneider’s case and that this impairment is what explains his inability to
perform abstract movements. I suggested earlier that these two claims may be reconciled if we consider that for Merleau-Ponty motor intentionality has both a 'centripetal' dimension, where it can operate independently of more abstract forms of intentionality and a 'centrifugal dimension' where it functions as a bridge between abstract, cognitive forms of intentionality (e.g., thoughts about movement) and actual movements. We can then understand him as claiming that, in Schneider's case, the centripetal dimension of motor intentionality is preserved, while its centrifugal dimension is impaired, the threat of inconsistency might be avoided.

But his means that an account of motor intentionality should aim at elucidating not just what distinguishes motor intentionality from more cognitive forms of intentionality but also how motor intentionality relates to these more cognitive forms of intentionality.

As Jensen (2009) points out, Merleau-Ponty's argument from abstract movement targets intellectualist models of action, according to which intentional bodily actions can be analyzed in terms of two independent components, a conscious intention representing the goal of the action and possibly the movements to be performed and the physical movements themselves caused by the intentions. Schneider's capacity to perform physical movements is intact, and he can form representations of abstract movements, such as drawing a circle in the air, since he can recognize when the movements he makes happen to be circular, yet he can not perform abstract movements in the normal way. Schneider's inability to perform abstract movements shows that this analysis is unsatisfactory. What remains a mystery and, for Merleau-Ponty, is doomed to remain one as long as we stay within an intellectualist framework is "by what magical process the representation of a movement causes precisely that movement to be made by the body" (Merleau-Ponty 2002, p. 160, n. 94). We thus need to appeal to motor intentionality to make bodily agency intelligible.
What we have said about motor intentionality up to this point is not enough to dissolve the mystery. Motor intentionality, understood as a basic form of intentionality, may well explain how bodily movements can be exercises of agency, be purposive and imbued with meaning and this independently of more abstract, conceptual, representational forms of intentionality. But we still lack an explanation of how motor intentionality and more cognitive forms of intentionality can be integrated and how our motor behavior can be responsive to our intentions. Merleau-Ponty claims that "the normal function which makes abstract movement possible is one of 'projection'" (Merleau-Ponty 2002, p. 128). Unless we can explain how this projection operates and how motor and cognitive intentionality are integrated, we are left with a projection process that appears no less magical than the process by which, in intellectualist accounts, the representation of a movement causes precisely that movement to be made by the body.

Several recent attempts have been made to address this issue. If a full explanation of human agency as integrated rational and bodily agency needs to appeal to both propositional attitude states like beliefs, desires, and intentions qua propositional attitudes and motor representations, we need to explain how intentions and motor representations can be coordinated and pull in the same direction. This problem is what Butterfill and Sinigaglia (2014) call the interface problem.

Butterfill and Sinigaglia argue that intentions and motor representations have distinct but complementary roles in explaining the purposiveness of actions and have a distinct representational format adapted to the function they serve. Intentions, understood in the standard way, are propositional attitudes with a characteristic role in practical reasoning and as such are subject to norms of rationality. We need to appeal to intentions and related propositional attitudes if we are to account for human agency as
the agency of beings who do things for reasons. The main functions associated with motor representations involve selecting the motor patterns needed to perform an action in a given situation and guiding and controlling their execution. As we saw in the previous section, to serve these functions motor representations must have a proprietary representational format, distinct from the format of intentions. Butterfill & Sinigaglia characterize the interface problem as the problem of explaining how it is that intentions and motor representations, having as they do different representational formats, are able to coordinate such that the action outcomes that they specify “non-accidentally match”.

Several approaches to the interface problem may be considered. What Butterfill and Sinigaglia call the common cause approach proposes that intentions and motor representations coordinate in virtue of sharing a common cause that triggers them both. The idea here is that a sensory state of the agent (e.g., a perception of a coffee mug) or an environmental stimulus (e.g., a coffee mug) triggers both an intention and a motor representation with aligned contents relating to the grasping of the mug. An advantage of this solution is that the difference in formats between these two representations does not raise any difficulties, since it is not in virtue of a causal interaction between them that they align. However, as Butterfill and Sinigaglia note, this is unlikely to provide a full solution to the interface. Neither intentions nor motor representations are always triggered by environmental causes. Intentions are often the result of deliberation or planning, and motor representations are frequently keyed to intentions rather than stimuli in the environment or an agent’s sensory states.

Wayne Wu (2011, 2015) develops another approach that appeals to intention-guided attention. Wu takes himself to be solving a slightly different problem, namely what he calls the Many-Many Problem. This is the problem that an agent faces of
selecting, out of many potential target objects for action and out of many potential actions on a target object, a specific action on a specific target object. On Wu’s view, intentions help an agent solve the Many-Many Problem by serving as structural causes that constrain the space of possible solutions. They do so in two distinctive ways, both centrally involving the deployment of concepts in their content. First, intention-guided attention identifies the object or objects to be acted upon from among competing objects. Thus, if one’s intention deploys the concept of, say, FORK, this thereby directs attention to the appropriate object in the agent’s perceptual field. Second, they activate appropriate motor representations. For example, an intention to GRAB one’s fork will guide the agent in attending to the spatial properties of the fork in appropriate ways, and activate motor representations constitutive of grabbing.

The first part of Wu’s solution to the Many-Many problem may be seen as a sophisticated variant of the common cause approach, the intention is not caused by an object in the environment, but it directs attention to the relevant object which in turn triggers a motor representation. However, this is only a partial solution to the Many-Many problem, merely reducing it to a One-Many problem. Hence, the second role assigned to intentions, where the action concept deployed in the content of an intention activates a motor representation appropriate to this action. However, from the point of view of the interface problem, the second part of Wu’s solution is problematic, as it appears to presuppose the existence of a connection between action concepts and motor representations rather than explaining it.

According to Butterfill and Sinigaglia (2014), the solution to the interface problem involves recognizing that the contents of intentions can be partially determined by the contents of motor representations and explaining what form this content-determining relation takes. Their explanation appeals to demonstrative and deferential
action concepts: the idea is that our intentions sometimes deploy demonstrative concepts that defer to motor representations specifying certain action outcomes, and thereby refer to those action outcomes, without any need for translation. Thus, on this proposal, we can consider the content of an intention to be “Do that!” and the demonstrative ‘that’ would defer to a motor representation referring to the relevant action. As Butterfill and Sinigaglia put it: “These demonstrative concepts would be concepts of actions not of motor representations, but they would succeed in being concepts of actions by deferring to motor representations. For any such concept, it is a motor representation which ultimately determines what it is a concept of” (Butterfill & Sinigaglia 2014, p. 134).

Mylopoulos & Pacherie (2016) have pointed out several disanalogies between ordinary instances of demonstrative reference and Butterfill's and Sinigaglia's proposed demonstrative deferral in intention that raise important difficulties for their view. Mylopoulos and Pacherie develop an alternative solution to the interface problem. Like Butterfill's and Sinigaglia's approach, this solution recognizes that the intention concepts deployed in the contents of intention can be partially determined by the contents of motor representations. However, they explain this content-determining relation by appealing to the notions of executable action concepts and motor schemas rather than to demonstrative deference. They propose that in order to properly interface with motor representations, intentions must have as constituents executable action concepts, where to have an executable concept for a given type of action one must have a motor schema for actions of that type. Motor schemas are more abstract and enduring representations than motor representations. They store knowledge about the invariant aspects and the general form of an action and are implicated in the production and control of action. On the one hand, they can be acquired through processes of
probabilistic inductive generalization from motor representations or from already extant schemas. On the other hand, the activation of a motor schema once learned will yield a motor representation, when the information needed to specify its parameters is provided, typically via attentional processes. Motor schemas would thus be what bridges the gap between intentions and motor representations, ensuring proper, content-preserving coordination between them.

6. Concluding remarks

An account of motor intentionality should aim at elucidating not just what distinguishes motor intentionality from more cognitive forms of intentionality but also how motor intentionality relates to these more cognitive forms of intentionality. While a wealth of conceptual and empirical work has helped sharpen our understanding of the distinctiveness of motor intentionality, our understanding of how motor and cognitive intentionality are integrated remains much more tentative, despite some recent attempts to address this issue. It remains debated as well whether, as Merleau-Ponty claimed, motor intentionality should be understood as non-representational or whether the notion of representation he worked with was too loaded and restrictive, opening the possibility that the contrast between cognitive and motor intentionality should be understood not as a contrast between representational and non-representational intentionality but rather as a contrast between conceptual and non-conceptual forms of intentionality. My own leanings are, as is probably already clear, towards the latter position. I favor a representational stance in part because, as I argued in section 4, motor 'representations' appear to meet sufficiently robust criteria for representationality, in part also (exhibiting here (again!) my own limitations and prejudices), because this
representational stance provides in my view a more promising starting point for understanding the interplay of motor and cognitive intentionality.
Acknowledgements

This work was supported by grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*

References


