

The phenomenology of action: A conceptual framework

Elisabeth Pacherie

▶ **To cite this version:**

Elisabeth Pacherie. The phenomenology of action: A conceptual framework. Cognition, 2008, 107 (1), pp.179-217. 10.1016/j.cognition.2007.09.003 . ijn_00352527

HAL Id: ijn_00352527
https://hal.science/ijn_00352527

Submitted on 13 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Available online at www.sciencedirect.com
ScienceDirect
COGNITION
Cognition 107 (2008) 179–217
www.elsevier.com/locate/COGNIT

The phenomenology of action: A conceptual framework ☆

Elisabeth Pacherie *

Institut Jean Nicod, ENS, EHESS, CNRS, Département d'Etudes Cognitives, Ecole Normale Supérieure, 29, rue d'Ulm, 75005 Paris, France

Received 31 July 2006; revised 3 September 2007; accepted 14 September 2007

Abstract

After a long period of neglect, the phenomenology of action has recently regained its place in the agenda of philosophers and scientists alike. The recent explosion of interest in the topic highlights its complexity. The purpose of this paper is to propose a conceptual framework allowing for a more precise characterization of the many facets of the phenomenology of agency, of how they are related and of their possible sources. The key assumption guiding this attempt is that the processes through which the phenomenology of action is generated and the processes involved in the specification and control of action are strongly interconnected. I argue in favor of a three-tiered dynamic model of intention, link it to an expanded version of the internal model theory of action control and specification, and use this theoretical framework to guide an analysis of the contents, possible sources and temporal course of complementary aspects of the phenomenology of action.
© 2007 Elsevier B.V. All rights reserved.

☆ Earlier versions of this paper have been presented at a workshop on the tempo of consciousness at Ecole Normale Supérieure in Paris in May 2005, at the Colloquium of the Institute of Cognitive Science in Lyons in January 2006 and at the Perception and Action Symposium held at Cornell University in May 2006. I am grateful to the many people who participated in the discussions in these occasions. I also thank Tim Bayne, Patrick Haggard, Joshua Knobe, Susanna Siegel, and Frédérique de Vignemont for comments on earlier drafts and three anonymous referees for this journal for many helpful comments and suggestions.

* Tel.: +33 1 44 32 26 81; fax: +33 1 44 32 26 86.
E-mail address: pacherie@ens.fr

0010-0277/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2007.09.003

Author's personal copy

Keywords: Action specification mechanisms; Consciousness of action; Sense of agency; Sense of control; Intention

1. Introduction

Given its central role in everyday experience, it is surprising that the phenomenology of action received until very recently such scant attention from action theorists and theorists of consciousness. Things are starting to change, however. In particular, improving psychological and neuroscientific methods have recently made the phenomenology of action an object of empirical investigation. One of the earlier pioneers was certainly Libet, whose famous studies on the 'readiness potential' were interpreted by many, including Libet himself, as evidence in favor of a skeptical attitude towards conscious mental causation. More recently, Wegner's psychological experiments and his claim that the conscious will is an illusion also promoted what Bayne and Levy (2006) aptly call 'will-skepticism'. These attacks on the traditional view of the structure of agency and the role the experience of agency plays within this structure did much to reawaken the interest of philosophers in the phenomenology of action. At the same time, further empirical investigations aimed at probing in more detail the phenomenology of action and its disorders have started yielding a wealth of new data, suggesting that will-skepticism may rest in part on too simplistic a view of the phenomenology of agency.

This burgeoning literature on the content and sources of the phenomenology of action highlights its many facets. A non-exhaustive list of proposed distinctions includes awareness of a goal, awareness of an intention to act, awareness of initiation of action, awareness of movements, sense of activity, sense of mental effort, sense of physical effort, sense of control, experience of authorship, experience of intentionality, experience of purposiveness, experience of freedom, and experience of mental causation. Beyond this terminological profusion, it remains unclear how these various aspects of the phenomenology of action are related, to what extent they are dissociable, and whether some are more basic than others. It also remains unclear what their sources are and how exactly they relate to action specification and action control mechanisms.

The purpose of this paper is to propose a conceptual framework allowing for a more precise characterization of the many facets of the phenomenology of action, of their relations to one another, and of their possible sources. One key assumption guiding this attempt is that the processes through which the component elements of the phenomenology of action are generated and the processes involved in the specification and control of action are strongly interconnected. In Section 2, I describe and motivate a dynamic model of intentions and action specification. In Section 3, I discuss in more detail the connections between this conceptual model and the internal model approach to action specification and control. I hope to show that the dynamic theory of intentions I propose provides a framework for thinking about

Author's personal copy

action that is both conceptually and empirically motivated. In Section 4, I provide a preliminary regimentation of the various components of the phenomenology of action. Sections 5 and 6 explore the extent to which these components may correlate with and possibly have their source in different aspects and stages of the processes of action specification and control. Section 5 focuses on our awareness of the content of our actions, i.e., our awareness of their goals and of the means employed to achieve these goals, while Section 6 focuses on the sense of agency for an action, i.e., the sense an agent has that he or she is the author of this action, and on the component experiences that contribute to the sense of agency.

2. A dynamic theory of intentions

The causal theory of action is the view that behavior qualifies as action just in case it has a certain sort of psychological cause or involves a certain sort of psychological causal process. In the last decades, this approach has gained wide currency. Yet, versions of causalism can take widely different forms depending on (1) what they take the elements of the action-relevant causal sequence to be and (2) what part of the sequence they identify as the action. With respect to (1), many philosophers have argued that in order to overcome some of the difficulties and shortcomings that plagued early belief/desire versions of the causal theory of actions, intentions – conceived as distinctive, *sui generis*, mental states with their own complex and proprietary functional roles – should be viewed as crucial elements of the action-relevant causal sequence. With respect to (2), several philosophers have also argued that actions should not be identified with mere bodily movements but with a larger chunk of the causal sequence. The dynamic theory of intentions proposed here shares both these insights.

Some of the functions attributed to intentions are typically played in the period between the initial formation of the intention and the initiation of the action. In contrast, other functions – in particular, their role in guiding and monitoring the action – are played in the period between the initiation of the action and its completion. Attention to these differences has led a number of philosophers to develop dual-intention theories of action. For instance, Searle (1983) distinguishes between prior intentions and intentions-in-action, Bratman (1987) between future-directed and present-directed intentions, Brand (1984) between prospective and immediate intentions, and Mele (1992) between distal and proximal intentions.[1]

Two, often implicit, assumptions of these dual-intention theories are problematic. First, they tend to assume that the role of the first of these two intentions is over once the second is in place. Their second, related, assumption is that action guidance and monitoring are the sole responsibility of the second intention. In contrast, I shall argue that we should distinguish three main stages in the process of action specifica-

[1] For a more detailed analysis of the difficulties and shortcomings of the belief/desire versions of CTA and for an evaluation of some of the proposals made by dual-intention theorists (see Pacherie, 2000, 2003, 2006).

Author's personal copy

tion, each corresponding to a different level of intention and each level of intention having a distinctive role to play in the guidance and monitoring of the action.

In the remainder of this section, I will therefore try and motivate a threefold distinction among intentions: distal intentions, proximal intentions, and motor intentions (D-intentions, P-intentions, and M-intentions for short).[2] This threefold distinction is based on an analysis of their different and complementary functional roles, of the different types of contents they involve and of their respective temporal scales. It is also informed by recent neuroscientific results. I will offer a characterization of the dynamics involved in the unfolding of intentions. Two levels of dynamics should be considered: the local dynamics specific to each level of intention and the global dynamics involved in the transitions from one level of intention to the next. With respect to the global dynamics of intentions, I shall argue that these transitions involve the construction of progressively more detailed representations of the action to be performed. Many aspects of the action that are initially left indeterminate are specified at the level of the P-intention and get further specified at the level of the M-intention. With respect to the local dynamics, it is also useful to distinguish, for intentions at each level, two phases of their internal dynamics: the upstream dynamics that culminate in the formation of the intention and the downstream dynamics manifested once the intention has been formed.

2.1. Distal intentions (D-intentions)

My notion of D-intentions is very close in certain respects to Bratman's notion of future-directed intentions (Bratman, 1987) Following his lead, we may stress three functions of D-intentions: as terminators of practical reasoning about ends, prompters of practical reasoning about means and plans, and intra- and interpersonal coordinators. The upstream dynamics of D-intentions – the dynamics of decision-making that lead to the formation of an intention – can be seen as involving two stages. The first stage is associated with the first of these three functions and involves deciding which end to pursue. The second stage in the upstream dynamics of D-intentions is linked to their functions as prompters of practical reasoning about means and as intra- and interpersonal coordinators. This reasoning must be internally, externally, and globally consistent. The various elements that form the building blocks of an action plan must be mutually consistent (*internal consistency*). The plan as a whole should be consistent with the agent's beliefs about the world (*external consistency*). Finally the plan must take into account the wider framework of activities and projects in which the agent is also involved and be coordinated with them in a more global plan (*global consistency*).

It is important to note that the distinction between the two stages in the upstream dynamics of intentions does not amount to a complete separation of the two pro-

[2] I have chosen to borrow Mele's terminology instead of those used by Bratman, Searle or Brand, as less likely to mislead one into thinking that the difference between distal and proximal actions is essentially or solely one of time and that distal intentions play no role in the guidance and control of actions once started.

Author's personal copy

cesses. Choosing among alternative ends is not just a matter of assessing and weighting their respective desirability; considerations of feasibility also play a role. A highly desirable end we have no means to attain or think we are highly unlikely to achieve may well be discarded in favor of a less desirable but more accessible goal. Yet typically, if a goal appears feasible, one can settle for it before engaging in a phase of detailed practical reasoning about the means to be employed to achieve it.

The upstream dynamics of D-intentions, although not completely context-free, are not strongly dependent on the particular situation in which the agent finds himself when he forms the D-intention or reasons from it. I can form a D-intention to act an hour from now, next week, two years from now or once I retire from work. This temporal flexibility makes it possible for an agent to form a D-intention to perform an action of a given type even though his present situation is not such as to allow its immediate performance. A D-intention is therefore in principle detachable from the agent's current situation and is indeed commonly detached from it. To that extent, its content is at least in part conceptual and descriptive. Of course, the current situation may sometimes prompt the formation of D-intentions, in which case, the content of the D-intention may involve indexical elements. Suppose someone kindly puts a plate of chocolate cake on my desk, I may well form the D-intention to eat this cake as soon as I finish writing this section.

The downstream dynamics of D-intentions is concerned with the high-level rational guidance and monitoring of the action. First, of course, the D-intention must be kept alive in prospective memory until the time comes to carry it out. When it does, one essential function of a D-intention is to ensure the rational control of the ongoing action. It is important to emphasize that what is specifically at stake here is the rational control of the action, since, as we shall see, P-intentions and M-intentions also have control functions, although of different types.

What should we understand rational control to be? Here, I will follow Buekens, Maesen, and Vanmechelen (2001) who describe rational control as taking two forms, 'tracking control' and 'collateral control', the second of which is often ignored in the literature. Tracking control involves making sure that each successive step in the action plan is successfully implemented before moving to the next step. It also involves revising the action plan when unforeseen circumstances make it impossible to successfully proceed as originally thought. Collateral control involves controlling for the side effects of accomplishing an action. The agent may notice undesirable side effects of her ongoing action and correct her way of accomplishing it in order to minimize them, or even abort the action. The main purpose of control as it is exercised at the level of D-intentions is to insure that the way the action is carried out does not flout the reasons the agent had for her action in the first place or violate the values, norms of coherence, general policies and rules of conduct to which she subscribes.

Although I have emphasized the rationality constraints at play at both the upstream and downstream stages of D-intentions, some qualifications are in order. We can distinguish between a *normative* and a *constitutive* conception of these rationality constraints. On a normative reading, rationality constraints apply to D-intentions in the sense that we ought to want to form D-intentions to do the things we think are best and we ought to want to have plans that are both internally and exter-

Author's personal copy

nally consistent insofar as this is a condition of their being successfully carried out in a non-accidental fashion. On a constitutive reading, an intention does not qualify as a D-intention unless these constraints are satisfied. It is clear that the rationality constraints, understood constitutively, are implausibly strong to say the least. Weakness of the will, emotional impingement on decision processes and the partial encapsulation of belief are commonplace phenomena. As social psychologists (and those we live with) never tire of telling us, most of us fall short of satisfying or even approximating the requirements of global and long-term consistency. If the rationality constraints are understood normatively, as I suggest they should, they can tell us only how intentions and plans *ought* to be formed for us to qualify as rational agents, while allowing that our cognitive behavior does not always meet the normative standards that it ought to.

Yet, insofar as "ought" imply "can", the rationality constraints that bear on both the upstream and downstream dynamics of D-intentions require the presence of a network of inferential relations among intentions, beliefs, and desires. Concepts are the inferentially relevant constituents of intentional states. 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, what follows is that for D-intentions to ever be such as to satisfy the rationality constraints they ought to, they must have conceptual content.

In a nutshell then, the content of D-intentions is both conceptual and at least partly descriptive. Because many aspects of an intended action will depend on yet unknown or unpredictable features of the situation in which it is eventually carried out, the initial description of the type of action leaves indeterminate many aspects of the action. A D-intention therefore always presupposes some measure of faith on the part of the agent, in the sense that the agent must trust herself at least implicitly to be able, once the time to act comes, to adjust her action plan to the situation at hand.

2.2. Proximal intentions (P-intentions)

As we did with D-intentions, we can distinguish between the upstream and the downstream dynamics of P-intentions. Their upstream dynamics are concerned with the generation of an intention to start acting now. The downstream dynamics concern the period that goes from the initiation of the action up to its completion.[3]

A P-intention often 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 to be performed, that is a representation that fits the

[3] Note that what I mean here by completion is simply the end of the action process, whether the action is successful or not.

Author's personal copy

specification inherited from the D-intention while anchoring it to the situation at hand. This anchoring process may be further characterized using the intentional schema theory proposed by Barresi and Moore (1996). Schemas are structures that organize information from different sources according to rules of integration to yield new perceptual or conceptual units. Schemas can be arranged in a hierarchy of levels according to their degree of dependence on current perceptual information.[4] The problem at the level of P-intentions consists in integrating 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. In other words, it consists in moving from an abstract schema (D-intention) to a schema that meets the brief set by the D-intention but is also constrained by current perceptual information.

At the downstream stage, P-intentions have to ensure that the imagined actions become current through situational control of their unfolding. As we did for D-intentions, we can distinguish between tracking and collateral control, where tracking control enables an agent to keep track of her way of accomplishing an action and to adjust what she does to maximize her chances of success, while collateral control is concerned with the side effects of accomplishing an action. Here, the main difference between P-intentions and D-intentions is that the former exercise tracking and collateral control with regard to the immediate goal and the situation as currently perceived, whereas the latter are concerned with the overall goal and the respect of global consistency and coherence constraints.

Insofar as a P-intention is tied to a corresponding ongoing action and gets deployed concurrently with it, it is subject to severe temporal constraints. These temporal constraints are of two kinds: cognitive and action-related. Firstly, P-intentions are concerned with aspects of the situation of action and of the activity of the agent that are explicitly perceived and conceptualized to some degree. They integrate a broad range of information about the agent, the goal and the context of action. This integration requires a rather long processing time. Therefore, the time scale of P-intentions is the time scale of explicit perception and thought. Second, their temporality is also constrained by what we may call the tempo of the action. This is literally the case when, say, one is playing a piece of music on the piano and must respect the tempo and rhythm of the piece. It is also the case in many other kinds of actions. In a game of tennis, one is allowed very little time to decide on how to return a serve. A slow tempo offers better conditions for online conscious guidance and control of the action, since the agent has more time to decide on adjustments or to consider and evaluate possible side effects. In contrast, when the tempo is extremely fast, possibilities for online conscious control may be very limited.

[4] Note that in their paper Barresi and Moore are mostly interested in intentional schemas underlying social understanding, that is, schemas integrating first and third person sources of information about object-directed activities, and in how, with the development of the imagination, these schemas can relinquish their dependence on current information. In contrast, I am interested in how we can move from abstract schemas to schemas constrained by current perceptual information.

Author's personal copy

2.3. Motor intentions (M-intentions)

D-intentions and P-intentions are responsible for high-level forms of guidance and monitoring, applying to aspects of the situation of action and of the activity of the agent that are perceived or conceptualized. However, work in the cognitive neuroscience of action shows that there also exist levels of guidance and control of an ongoing action that are much more specific, responsible for the precision and smoothness of its execution, and operate at a finer time scale.

M-intentions involve what neuroscientists call motor representations. I will not here attempt to review the already considerable and fast growing empirical literature on motor representations. For my present purpose, a brief description of the main characteristics of these representations will suffice. It is now generally agreed that there exist two visual systems, dedicated respectively to vision for action and for the identification and recognition of objects and scenes.[5] The vision for action system extracts from visual stimuli information about the properties of objects and situations that is relevant to action, and uses this information to build motor representations used in effecting rapid visuo-motor transformations. The motor representations produced by this system have three important characteristics. First, the attributes of objects and situations are represented in a format useful for the immediate selection of appropriate motor patterns. For instance, if one wants to grab an object, its spatial position will be represented in terms of the movements needed to reach for it and its shape and size in terms of the type of hand grip it affords. Second, these representations of the movements to be effected reflect an implicit knowledge of biomechanical constraints and the kinematic and dynamic rules governing the motor system. Thus, for instance, the movements of the effectors will be programmed so as to avoid awkward or uncomfortable limb positions and to minimize the time spent in extreme joint angles. 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 intrinsic characteristics (its shape and size) but also of the subsequent use one wants to make of it. The same cup will be seized in different ways depending on whether one wants to carry it to one's lips or to put it upside down. A given situation usually affords more than just one possibility for action and can therefore be pragmatically organized in many different ways. Recent work suggests that the affordances of an object or situation are automatically detected even in the absence of any intention to act. These affordances automatically prepotentiate corresponding motor programs (Grèzes & Decety, 2002; Grèzes, Tucker, Armony, Ellis, & Passingham, 2003; Tucker & Ellis, 1998).

One can therefore also distinguish two moments in the dynamics of M-intentions. The upstream dynamics lead to the selection of one among the typically several prepotentialized motor programs. When a M-intention is governed by a P-intention and

[5] This of course does not mean that the two systems do not interact in some ways. See Milner and Goodale (1995), Rossetti and Revonsuo (2000), and Jacob and Jeannerod (2003) for discussions of these interactions.

Author's personal copy

inherits its goal from it, the presence of the goal tends to increase the salience of one of these possible pragmatic organizations of the situation and thus allow for the selection of the corresponding motor program. Forming a P-intention to act on an object, say reach for a pen, typically involves focusing one's attention on the object that is to be the target of the action. As Campbell (2002) points out, we need to explain how the motor system manages to connect to the very object identified on the basis of perceptual experience. Campbell proposes that "conscious attention to the object will include some awareness of the location of the object, and that the target for processing by the visuo-motor system can be identified as 'the object at that location'" (2002, p. 55). In other words, common location is the binding principle for perception and for action, for objects represented at the level of P-intentions and objects represented at the level of M-intentions. The role of conscious attention is to define a target for the visuo-motor system. Once the target is defined, it is the job of the visuo-motor system to set the parameters for action on the object. Yet, it can also be the case that M-intentions are formed in the absence of a P-intention. In such cases, the upstream dynamics work in a different way. According to the model proposed by Shallice (1988) there is then a competition among motor programs, with the program showing the strongest activation being triggered as a result of a process he calls contention scheduling. The guidance and monitoring functions of M-intentions are exercised as part of their downstream dynamics. They are responsible for setting the precise parameters of motor commands and for fine motor adjustments and rapid corrections during execution. In Section 3, I will say more about the mechanisms that allow them to play these roles.

The motor system responsible for the production of M-intentions exhibits some of the features considered by Fodor (1983) as characteristic of modular systems. It is informationally encapsulated to some degree and its cognitive penetrability is limited. One illustration of its informational encapsulation is the fact that specification of motor commands by the motor system is, initially at least, relatively insensitive to perceptual illusions, both visual and tactile (Bridgeman, 1989; Haffenden & Goodale, 1998; Marcel, 2003). Yet, obviously the motor system cannot be fully encapsulated. If it were, it would be impossible to explain how a P-intention can trigger motor behavior, or how the way we grasp an object depends not just on immediate sensory affordances but also on our prior knowledge of the function of this object. But it is likely that the motor system has only limited access to information from other cognitive systems, including systems that underlie conscious perception.

A second feature of the motor system is its limited cognitive penetrability. Some global aspects of its operation appear to be consciously accessible and to be reflected in conscious motor imagery (Decety & Michel, 1989; Decety, Jeannerod, Durozard, & Baverel, 1993; Decety et al., 1994; Jeannerod, 1994). Yet, we are not aware of the precise details of the motor commands that are used to generate our actions, or of the way immediate sensory information is used for the fine-tuning of those commands (Fourneret & Jeannerod, 1998). For instance, several pointing experiments (Castiello, Paulignan, & Jeannerod, 1991; Goodale, Pélisson, & Prablanc, 1986) have shown that in a task where subjects have to point with their finger a target, they can do so accurately even on trials where the target is suddenly displaced by several

Author's personal copy

degrees and they have to adjust their trajectories. Moreover, they can do so while remaining completely unaware both of the displacement of the target and of their own corrections. One series of experiments devised by Rossetti and his co-workers (Pisella, Arzi, & Rossetti, 1998) is especially instructive. In one condition, a green target was initially presented and subjects were requested to point at it at instructed rates. On some trials the visual target was altered at the time of movement onset. It could either jump to a new location, change color or both. Subjects were instructed to point to the new location when the target simply jumped, but to interrupt their ongoing movement when the target changed color or both changed color and jumped. The results showed that when the target changed both color and position in a time window of about 200–290 ms, the subjects would make very fast inflight movement corrections, pointing at the displaced target instead of interrupting their ongoing movement. According to the explanatory scheme proposed here, this experiment may be interpreted as showing that M-intentions have their own dynamics, not entirely under the control of P-intentions.

These kinds of experiments also illustrate the fact that P-intentions and M-intentions operate at different time scales. The type of control exercised by P-intentions is, as we have seen, based on explicit perceptual experience. Temporal constraints on perceptual processes set a minimal temporal threshold for perceptual information to become accessible. Rossetti et al.'s experiments illustrate the existence of at least a partial incompatibility between the temporal constraints the motor system must satisfy to perform smooth online corrections and adjustments of an action and the temporal constraints on perceptual awareness.

2.4. General dynamics of intentions

Some characteristics of the macro-level dynamics of the transition from D-intentions to P-intentions and M-intentions can easily be inferred from what was said of the local dynamics of each level of intention. In particular, intentions at each level were assigned a specific role in, respectively, the rational, situational and motor guidance and control of the action. This implies that a D-intention does not cease to exist and play a role once it has given rise to a corresponding P-intention and similarly a P-intention does not go away once the corresponding M-intention has been generated. Rather, all three levels of intentions coexist, each exerting its own form of control over the action.

Yet the relation between the intentions at the three levels is not merely one of coexistence. They form an intentional cascade, with D-intentions causally generating P-intentions and P-intentions causally generating in turn M-intentions. More specifically, a D-intention may be said to trigger P-level processes providing them with an action plan that may still be mostly descriptive and abstract. The task of the P-level processes is then to anchor this action plan in the current situation and come up with a more detailed representation of the action tailored to the situation at hand. As another way to put it, P-level processes take as input both a specification of the action to be performed provided by the D-intention and perceptual information on the situation, including information on the agent's own state, and produce as out-

Author's personal copy

put a more detailed indexical representation. Similarly, P-intentions trigger M-level processes providing them with a situated representation of the action, say a target at a certain location and a given action to be performed on that target. Their task is then to process information on that target and to come up with a sensori-motor representation of the action that sets its parameters. Fig. 1 gives a summary representation of this intentional cascade and its temporal course.

What I have said so far on the dynamics of the three levels of intentions in no way implies that all actions require the presence of the entire intentional cascade. Some decisions to act are made on the fly and do not warrant a distinction between a D-intention and a P-intention. The existence of automatic, spontaneous or routine actions suggests that it is not even always necessary that I form a P-intention in order to start acting. It is interesting to note that even when a given action happens to be controlled by a P-intention, it is not necessarily this P-intention that triggered the action. While a routine action unfolds, I can become aware of what I am doing and decide whether or not I should go on with the action. If I decide to carry on, the action that had been initially triggered by a M-intention is now also controlled by a P-intention. It should also be noted that P-intentions can have varying degrees of control on the unfolding of an action. Well-practiced actions require little online control by P-intentions. In contrast, novel or difficult actions are typically much more closely controlled by P-intentions.

Let me close this section with two interrelated sets of remarks in response to one possible objection. One may worry that what I call M-intentions occupy a rather special status in the conceptual framework presented here and do not really deserve to be called intentions. I agree that M-intentions lack some of the features traditionally associated with intentions – for instance, their contents are not propositional and we may not be aware of them or have only partial access to their content. Yet, they

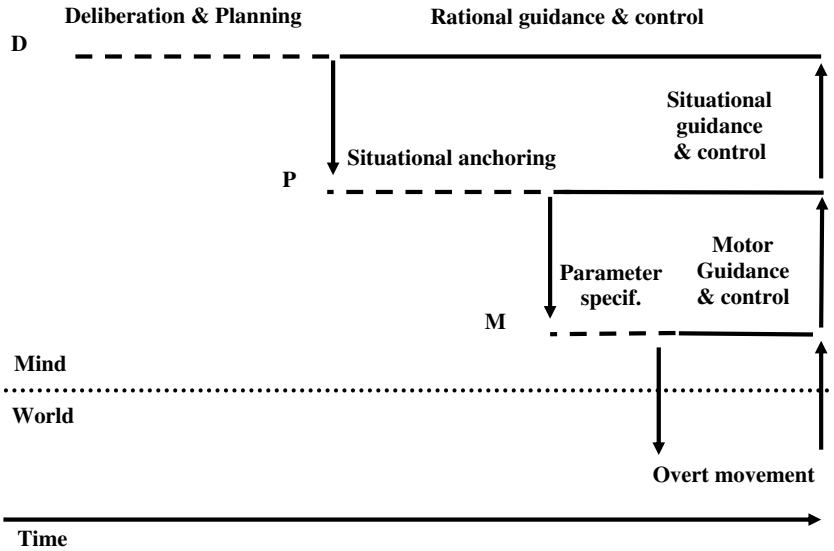


Fig. 1. The intentional cascade of D-intentions, P-intentions, and M-intentions.

Author's personal copy

share three core characteristics of 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 of fit 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. Insofar as they share these characteristics with D-intentions and P-intentions, one may, I think, have some justification in calling them intentions as well.

Perhaps another way of spelling out the objection is to say that intentions as traditionally conceived are directed at actions, whereas M-intentions are directed at mere movements. Many causal theorists, including Davidson (1980), have tended to identify actions with bodily movements. Indeed, early versions of the causal theory of actions were premised on the view that movements that are actions and movements that are mere happenings are indistinguishable in themselves and that the essential difference between them lies in their prior causal history. This view has been rightly criticized by several philosophers (Brand, 1984; Frankfurt, 1978; Hornsby, 1980; Searle, 1983). For instance, Searle argues for what may be termed a componential view of action.[6] An action is not a mere bodily movement, but consists in two parts, the movement and the intention-in-action that causes that movement. I would like to go further along this componential route, building on some insights of Harry Frankfurt. Frankfurt (1978) argues that what distinguishes an action from a mere bodily movement is the fact that the person is in some particular relation to the movements of his body during the time in which he's performing them and that this relation is one of guidance. He also claims that: "The fact that our movements when we are acting are purposive is not the effect of something we do. It is a characteristic of the operation at that time of the systems we are" (1978, p. 74) Frankfurt is right to point out that we cannot, on pain of infinite regress, conceive of our guidance of our movements while we are acting as requiring that we perform various further actions. His more cryptic claim that guidance is a characteristic of the operation of the systems we are can, I think, be more illuminatingly spelled out in terms of the notion of motor control. I therefore propose to say that an action in the minimal sense is an intentional movement, and consists of two parts: the bodily movement itself and the M-intention that causes and guides this movement. An intentional action in turn also consists in two parts: an action or intentional movement, understood in the sense just outlined, and the P-intention that causes and guides it. Finally, only D-intentions fit the traditional view of intentions as states characteristically directed at actions in the full-bloodied sense and as states distinct from the action they cause.

[6] Although when speaking of actions neuroscientists and psychologists often mean overt movements of the body, this not always the case. Jeannerod (2006), for instance, holds a componential view and argues that bodily movements are merely the overt part of actions that also necessarily involve a covert, representational part.

Author's personal copy

3. A theoretical model of action control

Central to recent work on motor control and motor learning is the idea that the motor control system makes use of internal models, which mimic aspects of the agent and of the external world.

The concept of internal models has been explored in depth by engineers who have proposed computational models incorporating the idea of control strategies based on internal models and have applied these models in the fields of robotics, neural networks and adaptive control. There is now growing evidence that similar strategies are used in human motor control (Frith, Blakemore, & Wolpert, 2000; Jeannerod, 1997; Jordan & Wolpert, 1999; Wolpert, 1997; Wolpert & Ghahramani, 2000; Wolpert, Ghahramani, & Jordan, 1995). There are two main kinds of internal models relevant for human motor control: forward and inverse models. Forward models (also called predictive models) mimic or represent the causal flow of a process in a system and use it to predict the next state of that system. Inverse models (or controllers) work in the opposite direction: they compute the motor commands that would have to be carried out to move a system from its current state to a certain desired state. Of special interest is the idea that the control of action depends in a large part on the coupling of inverse and forward models through a series of comparators, where comparators can be defined as mechanisms that compare two signals and use the result of the comparison for various kinds of regulation. Fig. 2 provides a typical illustration of the way these basic components are thought to be organized.

In a nutshell, inverse models compute the motor commands for achieving a desired state given the current state of the system and the current state of the environment. An efference copy of the motor commands is fed to a forward dynamic model that generates a prediction of the consequences of performing this motor command. This computational model also incorporates three kinds of comparators. A first kind of comparator takes as input representations of the desired state and of the predicted state. If a difference is found, an error signal is sent to the inverse model. Such a mechanism is useful in at least two ways. First, it can be used to maintain accurate performance in the presence of feedback delays. In most sensori-motor loops, the feedback delays are large and can result in inaccuracy or instability of the motor performance. The internal feedback of the predicted state of the system can be used before the actual sensory feedback is available to determine performance error and trigger corrections. This mechanism can also be used for mental practise and planning. During mental practise or in planning, forward models can be used to predict the sensory outcome of an action without actually performing the action. A comparison between desired state and predicted outcome could help the inverse model select between possible actions or it could improve the tuning of the inverse model.

A second kind of comparator mechanism compares the predicted consequences of a motor command with its actual consequences. The result of this comparison can be used to update the forward model and improve its functioning. It can also be used to filter sensory information and to distinguish the component that is due to self-movement from that due to changes in the world (Blakemore, Wolpert, & Frith, 1998,

Author's personal copy

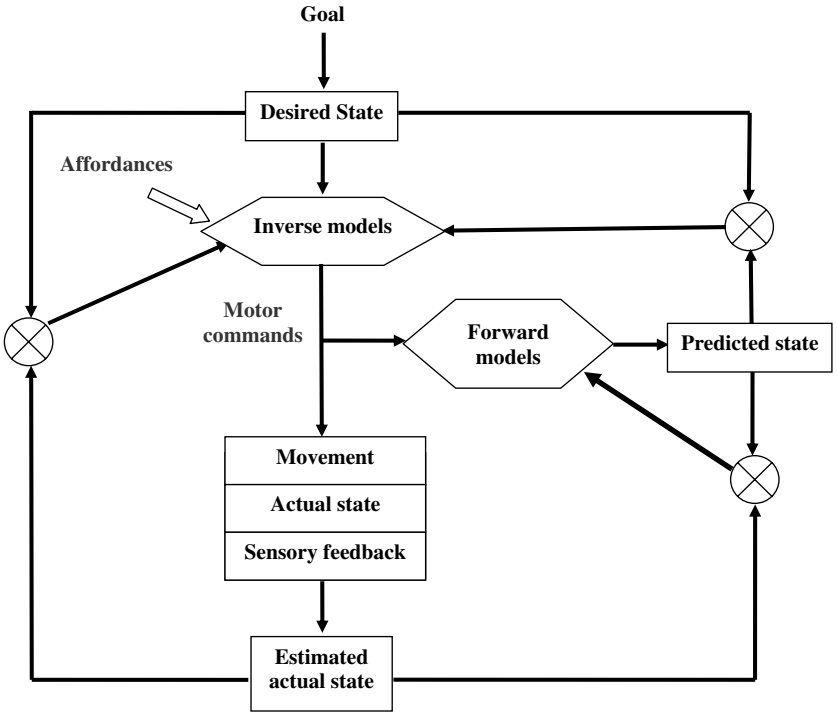


Fig. 2. The basic components of a motor control system based on internals models. Adapted from Frith et al. (2000).

1999). Finally, a third kind of comparison is between desired state and actual feedback. Errors derived from the difference between the desired state and the actual state can be used to update the inverse models and improve performance. This kind of comparison is therefore important for motor learning.

Although this way of thinking of the control of action has proven very fruitful, its main application has been to fine-grained aspects of motor control, such as the specification of motor commands and the prediction of the sensory consequences of their execution, corresponding to the level of M-intentions. Yet, there is no good reason why the idea of internal models should not be used in thinking about more global aspects of action specification. Presumably, the deliberative processes at work at the level of D-intentions make use of internal models of the world – both general theories such as folk-physics, folk-biology of folk-psychology and more specialized bodies of knowledge – as well as of the self-model the agent has of her desires, values, general policies and rules of conduct. Of course, the kinds of models exploited at this level have little to do with the internal models of the dynamics or kinematics of the motor apparatus. The contents represented at the level of D-intentions as well as the format in which these contents are represented and the computational processes that operate on them are obviously rather different from the contents, representational formats and computational processes operating at the level of M-intentions. Similarly, what are compared at this level are propositional (discrete) representations

Author's personal copy

of intended, predicted and actual states, rather than continuous, analogue representations requiring computationally more costly comparison mechanisms. Yet, the general idea that internal models divide into inverse models that given a goal compute means towards that goal and forward models which compute the consequences of implementing these means retains its validity at the level of D-intentions. And so does the idea that specifying an action plan and monitoring its execution rely on the coupling of inverse and forward models.

Similarly, it is highly plausible that action specification at the level of P-intentions makes use of internal models and that these internal models are different from both D-level and M-level internal models. On the one hand, as we have seen, the role of P-intentions is to anchor an action plan in a given situation of action and to select an appropriate action program. To play that role, they have to integrate a broad range of both conceptual and perceptual information about the current situation of the agent, the current goal and the context of action to yield a situated action plan, more specific than the typically rather abstract action plan formed at the level of D-intentions. On the other hand, as Rossetti's experiments discussed in the previous section illustrate, the representational resources available at the level of P-intentions are richer than the representational resources used by M-intentions and include information about conceptual or non-spatial perceptual properties of the situation not available to M-intentions. For instance, some of the confusion surrounding recent discussions of the exact nature of the impairment underlying delusions of control in patients with schizophrenia could probably be avoided by explicitly distinguishing between the predictions made by forward models at the level of P- and M-intentions.[7]

I therefore suggest that the information-processing model of action control in terms of internal models be explicitly combined with the threefold distinction among levels of intentions discussed in the previous section, thus yielding a richer theoretical framework for thinking about action. Fig. 3 provides a schematic representation of the more complex and, hopefully, more realistic view of action specification and control that results from this combination.

Let us take stock. What I have tried to do so far, is provide a characterization of the several stages in the process of action specification, from sets of beliefs and desires to detailed movements of several body parts. Extending an influential computational model of action control, I have also tried to offer a sketch of how the mind/brain could progressively generate, select, specify and monitor the parameters of intended actions, indicating when particular kinds of action representations are generated and how they are used. What I now want to explore is the idea that some of these information-processing events may have phenomenal counterparts and that links between conscious experiences during voluntary action and action specification processes may be identified by considering their respective contents and temporal properties.

[7] For a discussion, see Pacherie, Green, and Bayne (2006).

Author's personal copy

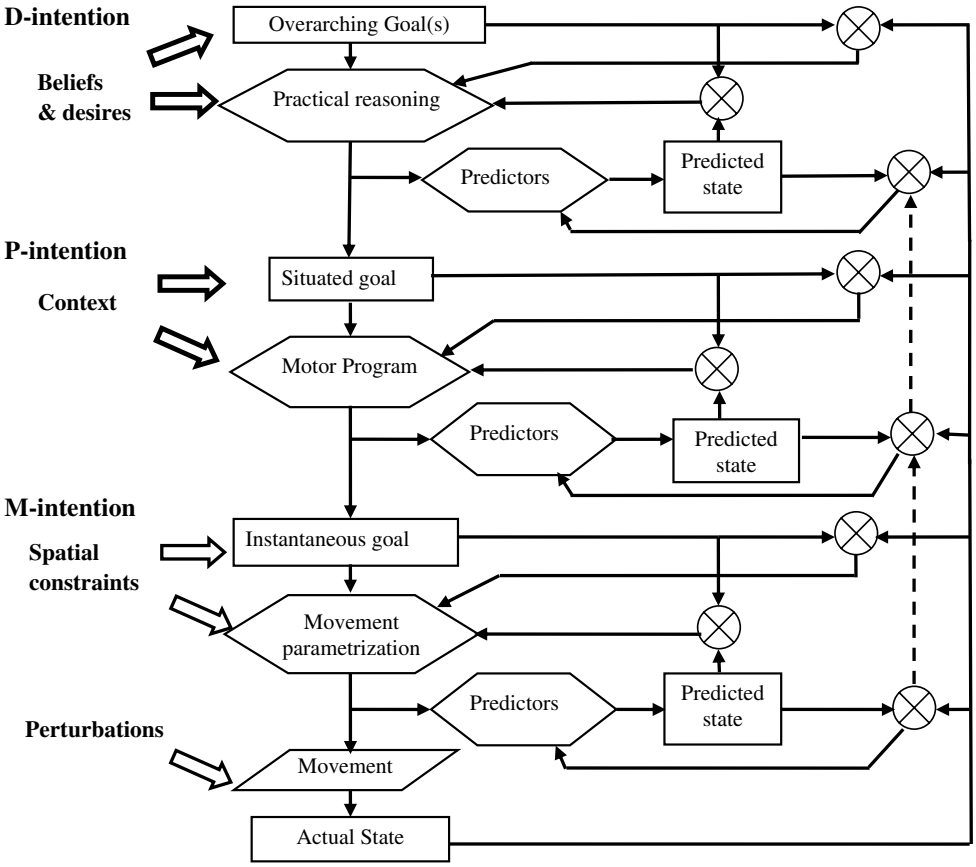


Fig. 3. A hierarchical model of action specification, with three levels of intentions and action control.

4. The phenomenology of action: a preliminary regimentation

It is now time to return to the issues with which this paper started. Both philosophical and empirical investigations highlight the fact that the phenomenology of action has many facets. This raises several questions: how are these various aspects of the phenomenology of action related? To what extent are they dissociable? Are some more basic than others? What are their sources and how exactly do they relate to action specification and action control mechanisms?

Let me start with some distinctions and a preliminary regimentation of these facets of the phenomenology of action based on what their content is about.

One distinction is between physical actions and mental actions and their respective phenomenology. Typically, physical actions involve the production of causal effects in the external world through movements of the body of the agent, while mental actions, such as attending to something or trying to remember the name of the person, do not. Here, I will focus on the phenomenology of physical actions, an important element of which is a sense of oneself as a physical agent producing physical effects in the world via its bodily interactions with it.

Author's personal copy

A second important distinction is between a long-term sense of agency and an occurrent sense of agency. The former may be thought to include both a sense of oneself as an agent apart from any particular action, i.e., a sense of one's capacity for action over time, and a form of self-narrative where one's past actions and projected future actions are given a general coherence and unified through a set of overarching goals, motivations, projects and general lines of conduct. The latter is the sense of agency one experiences at the time one is preparing or performing a particular action.

A third distinction is between detached and immersed awareness. Immersed awareness is the kind of non-reflective experience one has when one is fully engaged in an activity, while detached awareness requires a form of reflective consciousness, where the agent, so to speak, mentally steps back and observes himself acting or introspects what he is doing. Detached awareness can take at least two forms: a 'third-person' form where the detachment consists in the agent adopting the third-person stance of an external observer towards his own activity and a 'first-person' form where the agent introspects the thoughts and experiences he has while preparing and performing an action. In what follows, I'll be mostly concerned with what Marcel (2003), who draws similar distinctions, calls a minimal sense of agency, that is, our immersed awareness of our current action.

Yet, even so circumscribed, the phenomenology of acting is not something monolithic; it includes a number of distinguishable aspects. One way to draw these distinctions is in terms of the component elements of the content of our awareness of our current actions. First, some aspects of the phenomenology of agency concern the action itself, what is being done, while others concern the agent of the action, her awareness that she is acting or that she is the agent of the action. The former aspects, constituting what we may call awareness of action, themselves subdivide into what and how, i.e., awareness of the goal pursued and awareness of the means employed to attain this goal. The latter aspects of the phenomenology of action, the sense of agency proper, may itself be subdivided into a sense of intentionality or intentional causation, a sense of initiation and a sense of control. Note that this preliminary regimentation is not meant to preempt the question whether these various aspects are dissociable or not, for instance whether we can be aware of what we are doing independently of an awareness of how we are doing it or whether we can be aware of what we are doing without at the same time experiencing this action as ours.

Let us now move forward and examine how these component elements of the content of the phenomenology of agency could relate to component representations built at various stages of the process of action specification.

5. Awareness of action

5.1. What

Actions have a goal and typically the phenomenology of doing involves an element of purposiveness. In other words, we are aware to some degree that we are

Author's personal copy

engaged in purposive activity. According to the model of action specification described earlier, the goal of an action can be specified at the three levels of M-intentions, P-intentions, and D-intentions. What are their respective contributions to what-awareness – our awareness of what we are doing?

It is unclear whether we can be aware of our instantaneous motor goals, as they are implemented in a sensori-motor format. As I argued earlier, D-intentions, P-intentions, and M-intentions employ different representational formats and the representations at these levels have different temporal properties. There is some evidence that the representational format of M-intentions is incommensurable with the representational format or formats of our phenomenology and that M-intentions are too short-lived to be accessible to consciousness. One line of evidence in favor of incommensurability comes from the study of patients with visual agnosia. For instance, the patient DF, studied by Milner and Goodale (1995) was unable to recognize everyday objects, to visually identify simple shapes or to tell whether two visual shapes were the same or different. Yet her visuo-motor abilities were intact. When asked to pick up an object, she shaped her hand optimally for the grip, and when asked to post a card through a slit, she oriented the card correctly. The co-existence in such patients of impaired conscious visual perception and preserved visuo-motor abilities suggests that visuo-motor representations are normally not derived from conscious visual perceptions but built independently. It also suggests that conscious visual representations cannot be directly derived from intact sensori-motor representations. Similarly, in Rossetti's experiment described earlier, the surprise and frustration of the subjects at their inability to interrupt their ongoing movement when the target both changed color and jumped, suggests that they were unaware of the goal of their action as specified at the level of M-intentions. Yet, M-intentions may play a role in one's sense of agency for an action, as we'll see in Section 6.

In contrast to M-intentions, P-intentions specify our situated goals and represent them in a perceptual representational format readily accessible to consciousness. Through them we can be aware of what our immediate goals are. For instance, one might be aware of the movement of one's arm as an act of reaching for the doorknob, and perhaps as an act of reaching for the doorknob in order to open the door more widely than it is rather than to close it. Yet, when asked what I am doing, I will typically be able to provide a fuller answer than just 'I am opening the door'. For instance, I may say without hesitation that I am leaving my office to attend a seminar.

Can this be recovered from information available to P-processes? Arguably not. Different actions can realize the same further intention – I may have chosen to jump out of the window to go the seminar – and, conversely, the same action (to open a door) can realize a number of different further intentions. I may, say, open the door not to go to a seminar but to check the nature of a strange noise I heard in the corridor. Here's a more extreme example from Kevin Falvey: "Suppose a friend stops by my house and wants to go for a walk, and I say, "I can't; I am making bread." This could be true even if as I say it I'm sitting on the couch reading the newspaper – perhaps I am waiting for the bread to raise before putting it in the oven." (2000, p. 22). Such information about what one is doing can only be recovered from D-intentions. As Falvey's example shows forcefully, the immediate situation or one's imme-

Author's personal copy

diate actions may yield no clue as to what one is doing in this richer sense. Furthermore, because D-intentions can be the outcome of a conscious process of practical reasoning, they may also have access to one's reasons for acting. To the question "why are you making bread?", Falvey's character could answer, say, that he had guests coming for dinner and wanted to treat them with homemade bread.

In a nutshell, as Fig. 4 below illustrates, I suggest that what-awareness has two main sources. P-intentions provide us with a thin sense of what we are doing, restricted to our immediate goal, whereas D-intentions provide a thicker form of what-awareness, a representation of the kind of action it is at a more abstract level, and of one's reasons for performing it.

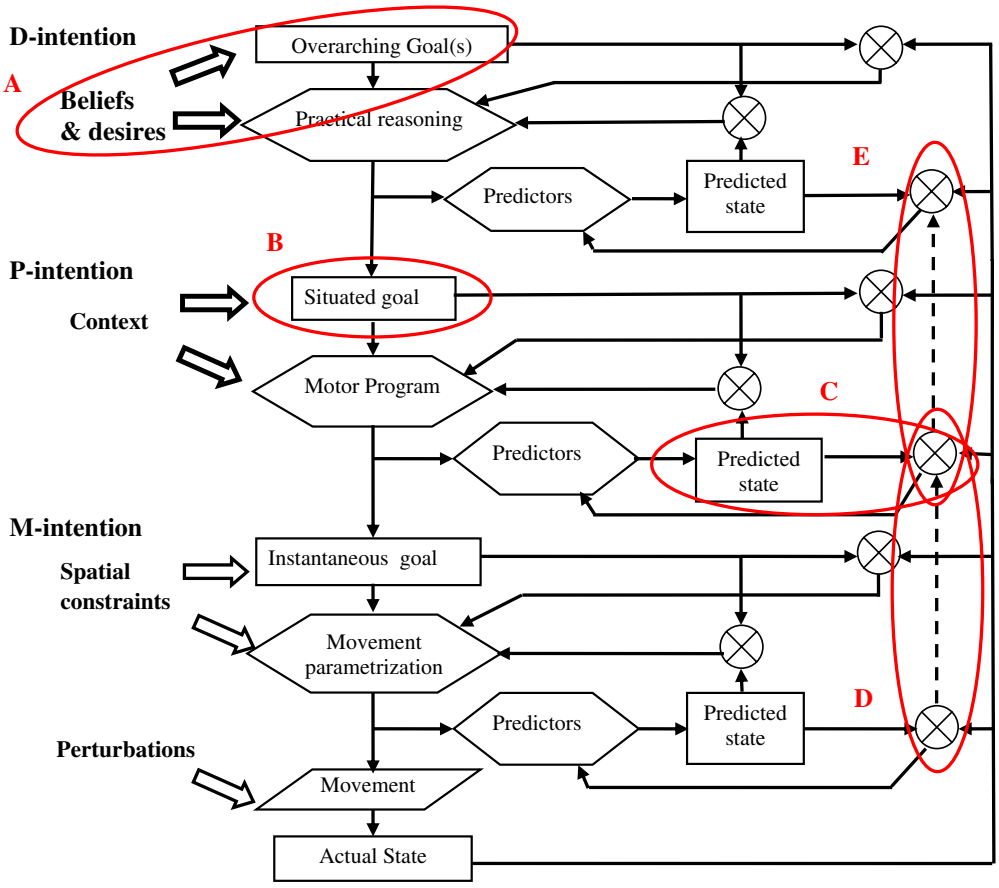


Fig. 4. The sources of what- and how-awareness. D-intentions provide a thick form of what-awareness, a representation of the kind of action it is and of one's reasons for performing it (A). P-intentions provide us with a thin sense of what we are doing, restricted to our immediate goal (B). Our awareness of our movements is normally limited and rests for the most part on our awareness of the predictions made at the level of P-intentions and on the comparison between these predictions and consciously available feedback (C). When the signals used to specify movement parameters and control execution at the level of M-intentions are too discrepant for errors to be automatically corrected, control is passed back to the level of P-intentions (D). When conscious compensation strategies at the level of by P-intentions fail, control is passed back to D-intentions for a more drastic revision of one's action plan (E).

Author's personal copy

5.2. How

Beyond being aware of what we are doing, in the sense of being aware of the goal of our action, we may also have some awareness of our specific manner of bringing about this desired result. Let us call this aspect of the awareness of action how-awareness. In the same way that in the process of action specification, there are three different stages of goal specification, there are also three different stages of means specification. Very schematically, at the more abstract level of D-intentions, means are typically represented as subgoals or subactions; at the level of P-intentions they are represented as movements of a certain type; finally, at the level of M-intentions they are represented as fully specified movements. Once again, we must ask which of these representations of means are accessible to consciousness and what they contribute to our how-awareness for our current actions.

But let me start with some remarks on the notions of means and of basic actions. First, as several commentators have noted following Goldman (1970), the phrase "doing A by doing B" does not always express a causal relation between the production of A and the production of B. At least three cases must be distinguished. The first is indeed causal generation. Doing A by doing B counts as an instance of causal generation when the production of B causes the production of A. For instance, turning on the light by flipping a switch or breaking the glass by dropping it on a hard surface count as instances of causal generation.[8] The second case is conventional generation, where doing B (in circumstances C) counts as doing A in virtue of a rule or convention that stipulates that is so counts. Thus, signaling a left turn by extending one's left arm or voting in favor of the motion by raising one's hand are instances of conventional generation. The third case is circumstantial generation, where doing B counts as doing A only if certain circumstances obtain. For instance, one breaks the world record for in the 100 m for men by running it in 9.77 s only in circumstances where no one has yet run this distance in 9.77 s or less. Such were the circumstances in June 2005 when Asafa Powell covered the distance in 9.77 s, but one would not break the world record today by running 100 m in 9.77 s.

Another important distinction is between two notions of basic actions. As Hornsby (1980) argued, one should distinguish between causal basicness and intentional basicness. Moving one's hand is causally more basic than turning on the light, because moving one's hand is what causes the light to go on.[9] An action is intention-

[8] Note that Goldman defends a fine-grained theory of action individuation according to which an action is an exemplification by an agent at a given time of an act-property, what he calls an act-token. Thus, for Goldman turning on the light and flipping a switch are two different actions since they are exemplifications of distinct act-properties. In contrast, coarse-grained theorists such as Davidson would claim that they are merely two different descriptions of the same action. Second and importantly, Goldman does not claim that my action of flipping on the switch *causes* my action of turning on the light but that it *causally generates* it. Causation and causal generation are for Goldman two different and mutually exclusive notions. He defines causal generation as follows: "Act-token A of agent S causally generates act-token A′ of agent S only if (a) A causes E, and (b) A′ consists in S's causing E." (1970, p. 23).

[9] Note that Hornsby (1980) actually introduces her two notions of basicness as relations between different descriptions of a single action rather than relations between different actions.

Author's personal copy

ally basic for one if one knows how to do it directly or "just like that" and can do it intentionally without having to do something else intentionally in order to do it. The notion of intentionally basic action is related to know-how and should be relativized to individual agents. Although for most people raising one's arm is an intentionally basic action, executing a trill may be a basic action for a professional pianist, but not for the novice piano player. Causal basicness and intentional basicness do not always go hand in hand. For instance, contracting such and such muscles is causally more basic than raising one's arm, but, unless perhaps one is a yogi, contracting such and such muscles is not something intentionally basic. Rather, raising one's arm is intentionally more basic than contracting such and such muscles; one intentionally contracts this set of muscles by intentionally raising one's arm.

My reasons for drawing attention to these distinctions will shortly become apparent. Returning to the issue of how-awareness, it is important to note that what counts as means vs. goal is level-dependent. What counts as means at the level of D-intentions are typically the subgoals in achieving the goal and these subgoals are represented in an abstract, semantic way. At this level, the relation of means to goals can take any of three forms introduced earlier: causal, conventional or circumstantial. When the D-intention is for a familiar action, say, going to work, we frequently do not need to explicitly consider the means to achieve this and as a result lack explicit awareness of the subgoals towards achieving our goal, even though they are in principle accessible to consciousness. Importantly, when moving from D-intentions to P-intentions and reaching a further, finer-grained, stage in the specification of action, these subgoals that are treated as means at the level of D-intentions become immediate situated goals while means correspond to ways of implementing these goals. From this level on, the question of means becomes the question of bodily movements and trajectories and means are to be understood in causal and circumstantial terms. So to what extent are we aware of our bodily movements? Awareness of movements appears to be modulated both by the intentional basicness of the action for the agent and by the degree of causal basicness of the action. Typically, agents have little or no awareness of how they accomplish actions that are intentionally basic for them as well as little awareness of the details of their movements beyond their more global parameters. For instance, I am normally aware that I am raising my arm to reach the cookie jar on the top shelf of the cupboard but have little or no awareness of how I raise my arm.

In a series of experiments, Jeannerod and co-workers (Fourneret & Jeannerod, 1998; Slachewsky et al., 2001) investigated subjects' awareness of their movements. Subjects were instructed to draw lines in the sagittal direction to a visual target with a stylus on a digital tablet. They could not see their hand; only the trajectory of the stylus was visible as a line on a computer screen, superimposed on the hand movement. A directional bias (to the right or to the left) was introduced electronically, so that in order to reach the target, the hand-held stylus had to be moved in a direction opposite to the bias. At the end of each trial, subjects were asked in which direction they thought their hand had moved by indicating the line corresponding to their estimated direction on a chart presenting lines oriented in different directions.

Author's personal copy

These experiments revealed several important points. Subjects accurately corrected for the bias in tracing a line that appeared visually to be directed to the target. When the bias was small, this resulted from an automatic adjustment of their hand movements in a direction opposite to the bias. Subjects tended to ignore the veridical trajectory of their hand in making a conscious judgment about its direction. Instead, they adhered to the direction seen on the screen and based their report on visual cues, thus ignoring non-visual (e.g., motor and proprioceptive) signals.[10] In other words, when biases remained small enough the visuo-motor system is able to appropriately use information for producing accurate corrections to reach a target, but that this information was not accessed consciously. However, when the bias exceeded a mean value of about 14°, subjects changed strategy and began to use conscious monitoring of their hand movement to correct for the bias and to reach the target. Yet, even though they consciously noticed the discrepancy between what they were doing and what they saw on the screen, subjects either experienced their movements as underestimates of their actual deviation or in the opposite direction to their actual adjusted movements. This transition from automatic to conscious control can be interpreted in at least two ways. According to Jeannerod and colleagues, when the discrepancy between the seen trajectory and the felt trajectory becomes too large to be automatically corrected, subjects become aware of it and use conscious compensation strategies. An alternative interpretation is that the change of strategy does not result from the conscious detection of a large discrepancy between visual and proprioceptive information but from the detection of a large discrepancy between predicted visual state and actual visual state. More precisely, the ongoing failure of the automatic control system to correct errors would result in control being passed back to the level of P-intentions in effect sending it a message of the form 'something is wrong' and conscious detection of error would then be the result of a comparison of predicted visual state and actual visual state at the level of P-intentions.

One reason for preferring the latter interpretation comes from the even more striking results of Marcel's vibro-tactile experiments (Marcel, 2003). By vibrating the biceps tendon at the elbow at certain frequencies, one can induce a reflex movement of the arm. If this movement of the arm is blocked, there occurs the illusion that the elbow is moving in the manner opposite to the reflex. Especially when the subject cannot see his stimulated arm, his hand feels to be in a position very different from its actual position. In his experiments, Marcel exploited this vibro-tactile illusion of limb position. In particular, in one condition subjects undergoing the illusion were asked to move their unseen hand to a target position indicated by a light. On some trials, the actual position of the arm was such that the agent had to move his arm to the left to reach the target position, while its felt illusory position suggested that the arm would have to be moved to the right to reach it. Subjects were asked (a) to draw with their

[10] The mode of response used in this experiment (indicating a line on a chart) may be thought to have induced a bias in leading the subjects to base their report on visual cues. Yet, in a variant of the experiment, subjects were asked instead to reproduce the movement they had made with their eyes closed. This change in the mode of response had no effect on their reports. Subjects still adhered to the direction seen on the computer screen (Jeannerod, personal communication).

Author's personal copy

other hand the movement they had to make to reach the target location (pre-movement drawing), (2) to move their unseen arm to the target location, (3) to draw with their free hand the movement they had just made (post-movement drawing). Quite interestingly, pre-movement drawing always followed the illusion, i.e., picturing the movement to be made as a displacement towards the right; the actual movement was always to the left showing no sensitivity to the illusion; in post-movement drawings, 60–70% of the subjects drew the movement in the same direction as their pre-movement drawing, suggesting that they had not noticed the difference between the movement they thought they had to make and the movement they had actually performed.

The fact that a majority of subjects in this experiment failed to notice this huge discrepancy between their predicted visual trajectory and their proprioceptive reafferences makes it doubtful that in Jeannerod's experiments the change of strategy of the subjects stemmed mainly from their noticing a much smaller discrepancy between vision and proprioception. An important difference between Jeannerod's and Marcel's experiments is that Marcel's subjects had no visual feedback, thus no way of comparing their visual predictions regarding the trajectory of their arm with visual reafferences and no way of noticing the discrepancy between the two kinds of signals. The nature of the tasks in both experiments may explain why proprioceptive feedback was neglected. In both cases, the action was directed at external goals in the form of visual targets. These actions where the primary aim is to achieve an external goal are to be contrasted with actions where the primary task is to make a movement. Experimental results from Wohlschläger, Engbert, Haggard, Clark, and Kalogeras (2003) suggest that for the former type of action our experience of acting is essentially outward-looking and dependent on information in exteroceptive modalities, both in the form of predictions and feedback, while for the latter type of actions proprioceptive information plays a crucial role.

In a nutshell, as illustrated in Fig. 4, our awareness of our movements rests for the most part on our awareness of the predictions made at the level of P-intentions and on the comparison between these predictions and consciously available exteroceptive feedback. When the action unfolds smoothly, this awareness is typically extremely limited. Action specification and action control mechanisms at the level of M-intentions operate automatically and remain outside the subject's subjective experience. When the signals this system uses to specify movement parameters and control execution are too discrepant for errors to be automatically corrected, failure becomes salient and control is passed back to the level of P-intentions. Thus, how-awareness typically becomes more vivid and more detailed when we are confronted with action errors too large to be automatically corrected. Yet sometimes, even conscious compensation strategies at the level of by P-intentions will fail. In such cases, control will have to be passed back to D-intentions for a more drastic revision of one's action plan.

6. Sense of agency

At first blush, it may appear strange to consider separately the question of awareness of action and the question of the sense of agency for our actions, where the sense

Author's personal copy

of agency is the sense that we are the author of that action. Philosophers often assume that that there is a constitutive link between the agent's awareness of an action and a sense of agency and hold a claim of immunity to error through misidentification for the self as agent. They assume either that our awareness of an action includes the agent of the action as part of its content or that the identity of the agent is guaranteed by the mode of access we have to the content of the action, where introspective as opposed to observational access to the content of an action would guarantee that the action is indeed ours. Yet, empirical evidence suggests that although awareness of action and sense of agency normally go together, they can sometimes come apart. The most striking illustrations are delusions of alien control in schizophrenia where a subject is aware of the content of the action she is executing but denies being the agent of this action. For instance, patients experiencing alien control will report:

> "My fingers pick up the pen, but I don't control them. What they do is nothing to do with me." (From Mellors, 1970, p. 18)

Or:

> "I felt like an automaton, guided by a female spirit who had entered me during it [an arm movement]." (From Spence et al., 1997).

Such dissociations between awareness of action and sense of agency can also occur in non-pathological conditions. Wegner's experiments, for instance, suggest that illusions of control – where we experience a sense of agency for actions someone else is doing – and illusions of action projections – where we de not experience a sense of agency for something we are doing – can be induced in normal subjects (Wegner, 2002).

These data suggest that to give an account of action awareness is not yet to give an account of the sense of agency. In this section, I will explore three main components of the sense of agency – the sense of intentional causation, the sense of initiation and the sense of control – each related in specific ways to predictor mechanisms.

6.1. The sense of intentional causation

Central to the sense of agency for an action is the sense that the agent is the cause of that action, what one may call the sense of intentional causation. The sense of intentional causation should be considered at two levels. At the more abstract level, one may feel that one's intention is the cause of one's action. At a lower level, one may feel that one's movement is causing some effect.

Both Humphrey (1992) and Wegner (2002, 2005) focus on the sense of intentional causation at the higher level. They argue that the sense of intentional causation[11] is

[11] Note that they use a different terminology. Humphrey uses the term 'sense of ownership' while Wegner speaks of 'the experience of conscious will'.

Author's personal copy

inferred from the existence of a match between a prior intention and an observed action.

> The experience of consciously willing our actions seems to arise primarily when we believe our thoughts have caused our actions. This happens when we have thoughts that occur just before the actions, when these thoughts are consistent with the actions, and when other potential causes of the actions are not present. [...] In essence, the theory suggests that we experience ourselves as agents who cause our actions when our minds provide us with previews of the actions that turn out to be accurate when we observe the actions that ensue. (Wegner, 2005, p. 23)

Some of Wegner's experiments suggest that the experience of intentional causation can be non-veridical, with either agents having an experience of agency for an action they have not actually caused (Wegner & Wheatley, 1999) or, conversely, attributing to others their own actions (Wegner, Fuller, & Sparrow, 2003). The capacity to predict or, as Wegner would say, to preview what will happen seems to play a key role in the experience of intentional causation at this level. Thus, Wegner and colleagues (Wegner, Sparrow, & Winerman, 2004) had participants watch themselves in a mirror while another person behind them, hidden from view, extended hands forward on each side where participants' hands would normally appear and performed a series of movements. When participants could hear instructions previewing the movements, their sense of agency for these movements was enhanced, but such vicarious agency was not felt when the instructions followed the movements.

Although the presence of a match between a prior intention and an observed action may certainly contribute to the sense of intentional causation for an action, there are two main problems with the view that the sense of agency arises solely or primarily when there is such a match. First, as several commentators (Marcel, 2003; Pacherie, 1996, 2001; Spence, 2001) have pointed out, prior intentions or awareness thereof do not seem to be necessary for the sense of agency. On many occasions, we cannot remember what our prior intentions were and yet do not disown the actions. Furthermore, many of our actions, impulsive, routine or automatic, are not preceded by conscious intentions and yet we own them. Second, awareness of an intention and of a match between intention and action does not seem sufficient for a sense of agency. For instance, Frith, whose early account of delusions of control in schizophrenia posited the existence of a deficit in intention-monitoring resulting in the *loss of awareness of 'willed' intentions* to act (Frith, 1992), later acknowledged that this "is inconsistent with the patients' ability to follow the commands of the experimenter, to avoid showing utilization behavior, and to correct errors on the basis of sensory feedback about limb positions (which requires comparisons of intended actions and their consequences)" (Frith et al., 2000, p. 1784). Similarly, as Nahmias (2005) points out, Wegner's three conditions of priority, consistency and exclusivity between conscious intentions and actions can be met, without one thereby experiencing oneself as the agent of the action.

Author's personal copy

The fact that one may experience a sense of intentional causation for one's acts in the absence of a corresponding conscious prior intention suggest that other types of matches than just the match between a prior intention and an observed action play a role. One such match is between an action and its consequences. Haggard and colleagues (Haggard & Clark, 2003; Haggard, Clark, & Kalogeras, 2002) have shown that when a voluntary act (a button press) causes an effect (a tone), the perceived time of initiating the act is closer to the perceived time of the effect. Specifically, the action (the button press) is shifted forwards in time towards the effect it produces, while the effect is shifted backwards in time towards the action that produces it.

Haggard calls this phenomenon intentional binding and suggests that it may be an implicit measure of the sense of agency. While acknowledging that the exact neural mechanisms that give rise to intentional binding have yet to be elucidated, Haggard suggests it probably derives from predictive mechanisms of action control. There are several lines of evidence for this view. First, intentional binding depends critically on the presence of voluntary movement and requires an efferent signal. When similar movements and auditory effects occur involuntarily, the binding effect is reversed and cause and effect are perceived as further apart in time than they actually are. Second, the experience of intentional causation appears to be constructed at the time of the action itself, as an immediate by-product of the motor control circuits that generate and control the physical movement itself. Haggard and Clark (2003) used TMS to insert occasional involuntary movements of the right finger at a time when the subject was intending to press the button, but had not yet done so. They found that if the intention was interrupted by an imposed involuntary movement causing the button press, followed by the tone, intentional binding did not occur. These results show that an (incomplete) intention followed by the appropriate effect is not sufficient for intentional binding.

Third, intentional binding requires reliable relations between actions and effects (Haggard et al., 2002; Sato & Yasuda, 2005). Thus, if a button press has been reliably followed in the past by a certain tone after a given temporal interval, intentional binding will be strong when the key press again produces the same tone with the same timing but reduced if the key press evokes a tone at a different pitch or with a different temporal interval. Yet, it is important to note that forward models are not fixed entities but must be learned and updated through experience. There is some evidence that intentional binding reflects this updating process. In the absence of a prior representation of an action-effect relation, the only evidence about the existence of such a relation comes from current sensory data and what may be inferred from these data. Such inferential processes are needed to uncover and learn contingencies and they may influence the sense of agency for action effects during the course of learning. In a recent experiment, Moore and Haggard (2007) investigated the respective roles of prediction and inference in intentional binding. They found that when the probability of the effect was high, the predictability of the effect influenced the perceived time of the action and that a delay in the awareness of the action indicative of binding occurred even on trials where the action produced no effect; in contrast, when the predictability was low, this delay was found only on trials where

Author's personal copy

the auditory effect occurred, reflecting the presence of an inferential process. In addition to these long-term effects, they also found that very recent reinforcement of the action-effect relation increased binding on action only trials.

Moore and Haggard speculate that these results are best explained using a Bayesian framework. This Bayesian framework would integrate prior knowledge, both long-term and short term, and new sensory data, weighting each source according to its reliability. Thus, when effect probability is high, there should be a strong prior representation of the action-effect-relation, which should make possible a clear prediction about the occurrence of the effect. In contrast, when the effect probability is low, there should be a weak prior representation of the action-effect relation, allowing no clear prediction about the occurrence of the effect. Within this Bayesian framework, the degree of intentional binding between action and effect and, correlatively, the degree to which one experiences a sense of intentional causation would be based on the comparison of *weighted* prediction and *weighted* feedback, rather than prediction and feedback *simpliciter* (see Fig. 5).

Although the intentional binding of an action and its effects may underlie the experience of intentionality or intentional causation for an action, it is unclear whether, as Haggard and Clark (2003) seem to suggest, an experience of intentional causation for an action is tantamount to an experience that *I* caused this effect and am therefore the author of that action. In a series of studies Wohlschläger et al. (2003) showed that this binding effect and the associated sense of intentional causation also occur when we observe other people's actions. Subjects had to estimate the onset time of button press that they executed themselves or that they observed being executed by someone else or else by a mechanical device. The estimate of the machine actions was always different from those of self- and other-generated actions, whereas the latter two were indistinguishable and showed a binding effect.

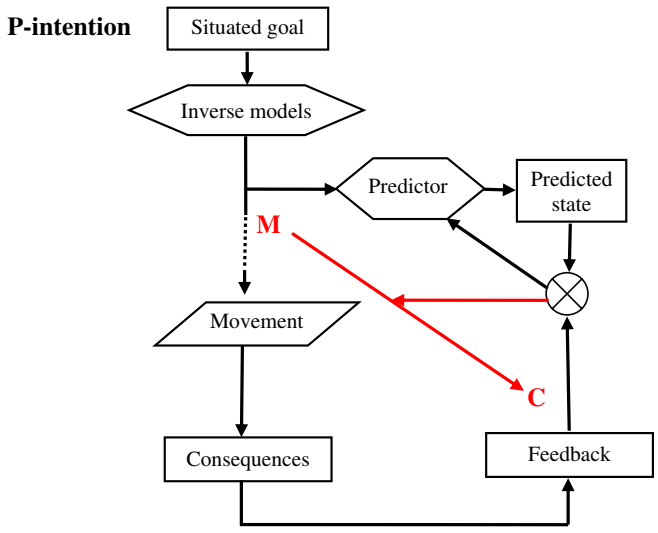


Fig. 5. The sense of intentional causation at the level of P-intentions. Weighted prediction and weighted feedback are compared using a Bayesian computation. A binding mechanism binds movement (M) and consequence (C) depending on the result of this computation.

Author's personal copy

These surprising findings need not be inconsistent with the predictive account of intentional binding favored by Haggard, provided one assumes that the predictive mechanisms used for action control also operate when one observes someone else acting.[12] If, however, intentional binding can occur for observed actions as well as self-produced actions, it cannot be the basis of the sense of authorship for an action. Frith (2005) points out that if the two kinds of experience are distinct, then it should be possible for one to be impaired while the other remains intact. Indeed he suggests that a patient with delusions of control may experience a strong sense of intentional causation for an action he has just performed while lacking a sense of authorship for that action.[13] This dissociation of the two experiences would form the basis of an explanation of delusions of control. It is unclear, however, whether the reverse dissociation could occur, with someone feeling a strong sense of authorship for an action in the absence of an experience of intentional causation. If indeed there are no dissociations of this latter type, the experience of intentional causation associated with intentional binding, although not sufficient to generate an experience of authorship for an action, would still be a necessary component of that experience.

6.2. The sense of initiation

According to Frith, Blakemore and Wolpert's revised account of delusions of alien control (Blakemore, Oakley, & Frith, 2003; Blakemore, Wolpert, & Frith, 2002; Frith, 2005; Frith et al., 2000), in normal circumstances when an agent is performing an action, she is aware of (i) her goal, (ii) her intention to move, (iii) her movement having occurred, and (iv) her having initiated her movement. In contrast, a patient with delusions of control has normal awareness of (i)–(iii) but not of (iv).[14] According to this model, awareness of initiating a movement depends on awareness of the predicted sensory consequences of the movement. This view is based on evidence that awareness of initiating a movement in healthy subjects is reported by the agent between 80–200 ms before the movement actually occurs (Libet, 1985; Libet, Gleason, Wright, & Pearl, 1983).

In experiments extending Libet's work, Haggard and colleagues (Haggard & Eimer, 1999; Haggard & Magno, 1999) confirmed that both intention judgments, corresponding to awareness of an intention to move, and movement judgments, corresponding to the awareness of movement onset, preceded actual movement, but

[12] There is increasing evidence that common brain circuits are used both to control one's actions and to represent observed actions performed by others (see Blakemore & Decety, 2001; Fogassi & Gallese, 2002; and Grèzes & Decety, 2001 for reviews). Computational frameworks showing how predictive models used for motor control can also be applied to others in social interaction, including action observation, have also been proposed (Wolpert, Doya, & Kawato, 2003).

[13] Note that although we exploit the same distinctions, Frith's terminological choices differ from mine. What I call the experience of intentional causation, he calls the experience of agency or intentionality; what I call the sense of agency or authorship, he calls the experience of ownership or of being in control of an action.

[14] For a more detailed discussion of Frith's revised model and its possible limitations, see Jeannerod and Pacherie (2004) and Pacherie et al. (2006).

Author's personal copy

they also showed that both types of judgments were unrelated to the general readiness potential but covaried with the lateralized readiness potential.[15] This suggests that awareness of an intention is tied not to the general aspects of action preparation but to the selection of a specific motor program. This also suggests that both awareness of intention and awareness of movements are associated with pre-motor processes rather than motor processes themselves. In another experiment using Transcranial Magnetic Stimulation (TMS), Haggard and Magno (1999) showed that applying TMS over the primary motor cortex created a large delay of the actual reaction time (movement onset) but a much smaller delay of the time of awareness of movement, whereas applying TMS over pre-motor areas, specifically the SMA, led to a much smaller delay of actual reaction time but to a greater delay in the awareness of movement. These data support the view that awareness of movement onset is generated upstream of the primary motor cortex but downstream of the pre-motor structures.

Interestingly, in another study, Sirigu and colleagues (Sirigu et al., 2004) showed that patients with parietal damage could report when they started moving but not when they first became aware of their intention to move, while cerebellar patients behaved like normal subjects. Both the parietal cortex and the cerebellum are thought to be involved in the predictive control of action: the cerebellum would make rapid predictions about the sensory consequences of self-generated movement at a very low level of movement execution, while the parietal lobe would be concerned with more high-level predictions and more cognitive aspects of action (Blakemore & Sirigu, 2003). (Within the present framework, we could say that the cerebellum is involved in predictive control at the level of M-intentions and the parietal lobe with predictive control at the level of P-intentions.)

These experimental results suggest that awareness of an intention to move is correlated with activity in the parietal cortex, corresponding to an early stage in motor preparation. We can speculate that awareness of an intention to move corresponds to the detection of a match between the predictions made by the parietal system and the desired outcome. We can also speculate that when we are preparing to make a movement, an inhibitory process is employed to prevent the movement from happening too early. The inhibition should be released only when a match is found between the predicted consequences of the movement and the desired outcome. Furthermore, when the prepared movement is part of a sequence of movements, inhibition should be released according to its position in the sequence and to the degree of completion of earlier elements of the sequence.

As Haggard (2003) points out, the co-existence of awareness of intention and awareness of movement onset within a single narrow window of pre-motor processing suggests that binding these two representations may be important. In particular, the efferent binding of these two representations may underlie the sense of initiation for the action, where the sense of initiation is not just the sense that we started mov-

[15] Intention judgments correspond to what Libet and Haggard call W-judgments, and movement judgments to what they call M-judgments. The intention one is aware of when making a W-judgment is what I would call a P-intention.

Author's personal copy

ing, but the sense that we started moving in accordance with our intention. It is important to note that the efferent binding at stake here is not the intentional binding discussed in Section 5.1. What I suggested earlier is that the intentional binding of an action (movement onset) and its consequences gives rise to the experience of intentional causation; what I am suggesting here is that the binding of intention and movement onset gives rise to the experience of initiation, as represented in Fig. 6.

One should further note that in daily life actions typically involve complex sequences of movements rather than the simple key presses or finger movements favored in experimental settings. Although it may well be that for simple actions the sense of initiation is an experience that arises only once at the outset of the action, we should not think that this is generally true. Think of a prosaic action such as getting a drink of water. It involves a number of sub-steps such as lifting the bottle, pouring water into a glass, putting down the bottle, reaching for the glass and bringing it to one's lips. To start acting one need not have already computed the motor commands for each of these steps. Rather, one computes the motor commands corresponding to the initial steps and continues computing the motor commands for the remaining steps while already carrying out the first steps. Yet, the motor commands for each step must be validated through a comparison of their predicted consequences with the desired outcome and inhibition released when appropriate. Proper binding of intention and movement onset may be crucial for the correct temporal execution of elements in the sequence. With actions involving a series of steps, the sense of initiation does not reduce to a single brief experience at the very beginning of the sequence. Rather, we have a series of events of initiation and a comparable series of experiences of initiation.

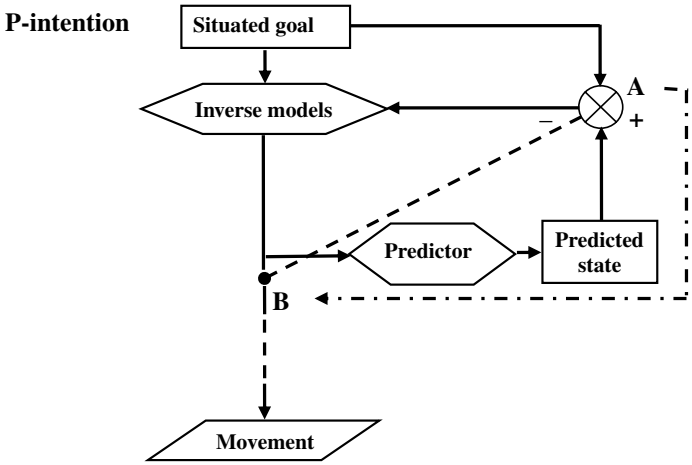


Fig. 6. The sense of initiation at the level of P-intentions. The dotted arrow with the blunt head represents an inhibitory link. The execution of the motor command is inhibited as long as there is no match between desired and predicted state at the level of the P-intention. The awareness of an intention to move (A) corresponds to the detection of such a match. As a result of this match inhibition is released, giving rise to awareness of movement onset (B). The sense of initiation would result from the efferent binding of A and B, represented here by the dash-dot.

Author's personal copy

6.3. The sense of control

Although the sense of initiation may be a crucial component of the sense of agency for an action, it does not seem to offer the guarantee that the ensuing action will be owned by the subject. For instance, we may sometime feel that we initiated an action but do not control its course. If something unforeseen happens, the action may get out of hand so to speak. We may feel we've lost control over it and this feeling may result in a reduction or even abolition of the sense of agency for the action. Similarly, it is unclear whether the experience of alien control in schizophrenia is always associated with a lack of sense of initiation for the action. In the patient's report from Spence et al. (1997) I quoted earlier, the patient described his experience as follows: "I felt like an automaton, guided by a female spirit who had entered me *during* it [an arm movement]" (my emphasis). This suggests that the experience of alien control for an action may sometime appear after the action has started.

The sense of control may therefore be another crucial contributor to the normal sense of agency for an action. Yet, it is doubtful whether the experience of control is itself a simple, elementary phenomenon. It seems rather that the sense of control can take different forms and varies along several dimensions and should be conceived as a compound of more basic, partly dissociable experiences. First, it should be noticed that talk of sense of control for an action can refer to two rather different kinds of experience. On the one hand, it may refer to the extent to which *one feels in control* of an action, where at one extreme everything happens exactly as expected and the agent feels in full control of his action and at the other everything goes astray and the agent feels completely powerless. On the other hand, by sense of control we may refer to the sense that *one has to exert control* to generate and maintain an appropriate action program despite perturbating factors. Normally control in this latter sense is felt as effortful: the more one has to exert control to attain one's goal, the more effortful the action feels.

Second, according to the model of action specification and control I described in Section 3, intentions at each of the three levels I distinguished exert their own specific form of control and guidance over the action. Control and guidance at the level of D-intentions must ensure that the successive steps in the action plan are implemented. They must also ensure that the adjustments and corrections that may have to be made while the action unfolds lead to the attainment of the overall goal and do not flout the rationality and consistency constraints on which the action was premised. P-intentions anchor action plans both in time and in the situation of action. Control and guidance at this level must take into account the characteristics of the agent, the target of action and the surrounding context at the time of action. Actions are represented in terms of motor programs and of the perceptual consequences of their execution. Guidance and control consist in ensuring that the motor program is implemented and, if necessary, in adjusting it so it yields the expected perceptual consequences. Finally, M-intentions specify the detailed parameters of the selected motor program. Motor control processes are responsible for the fast online adjustment and fine-tuning of these parameters. We may therefore speak of rational con-

Author's personal copy

trol at the level of D-intentions, situational control at the level of P-intentions and motor control at the level of M-intentions.

The experimental evidence we discussed in Section 5.2 suggests that motor control at the level of M-intentions is automatic and that typically subjects are not aware of the nature of the adjustments and corrections made at this level. It is only when the discrepancy between predicted and actual sensory consequences of the movements becomes too large to be automatically corrected that it becomes accessible to consciousness. This suggests that the sensori-motor signals used for motor control could contribute something to the experience of control. This contribution would, however, be limited in two ways: first, sensori-motor signals would only contribute to the experience of control when large discrepancies are present and, second, these signals would only modulate the extent to which one feels in control of the action but would not tell us the aspects in which the action escapes our full control.

At the level of P-intentions, the sense that one is in control would rely on the perceived match or mismatch between the predicted perceptual effects, corresponding to the situated goal, and the actual perceptual effects; while at the level of D-intention, it would depend on whether our action plan is carried out successfully, in other words on the conceptual match between the desired, predicted and actual consequences of its successive steps. It is important to note that these two different forms of the sense that one is in control can come apart. Perfect situational control over one's motor action will not guarantee the achievement of the overarching goal if the action plan was badly thought out in the first place and, conversely, one's general goal may be achieved despite approximate situational control.

At this point it is important to stress both the similarities and dissimilarities between the sense of intentional causation discussed in Section 6.1 and the sense of control. Both experiences rely on predictions made by internal models and on the comparison of these predictions with actual feedback. As we have seen, intentional binding between an action and its consequences can occur even when the agent had neither intended nor predicted these consequences, given a weak prior and reliable sensory feedback. In contrast, the sense that one is in control requires strong priors and feedback that matches both the predicted and desired states. Skilled motor behavior relies on accurate predictive models and practice is needed to learn such predictive models. A novice tennis player, say, has not yet learned the relevant contingencies and typically feels he has little control over what he is doing even when the consequences of his actions are as desired. Sense of control and sense of intentional causation therefore go hand in hand when the agent has a reliable internal model, hence strong priors, and both desired and predicted states are congruent with feedback, but they can dissociate in the absence of strong priors, where the agent may retain a sense of intentional causation while lacking a sense of control.

The sense that one is in control is often, I suggest, a compound of three more basic experiences: the sense of motor control, the sense of situational control and the sense of rational control. In all three cases, the degree to which one feels in control depends on a comparison between desired, predicted and actual states, where the better the match the stronger the sense of control. One important difference between motor control on the one hand and situational and rational control on the other is

Author's personal copy

that when one does not feel in full motor control one is simply aware that something is wrong, whereas when one's does not feel in full situational or rational control one can be aware not just that something is wrong but of what is wrong.

When something is wrong, one normally feels one has to exert control to keep the action on track. The type of control one has to exert depends on the nature of the perturbating factors. Perturbations may be due to external or to internal factors, may be physical or not, may have been anticipated or not, and may affect motor, situational or rational control. Depending on their nature, resistance to perturbating factors can require either physical or mental effort. For instance, lifting a heavy box requires physical effort, reading in a noisy environment requires concentration, and inhibiting a prepotent but inappropriate response requires mental effort. When the effect of perturbating factors has been anticipated (I know that the box is heavier than it looks, that I am in England and should drive on the left side of the road, that solving this problem is difficult and requires concentration), the amount of force or the attentional resources needed are pre-programmed and would be part of our awareness of the content of our intention immediately prior to action. But when these disturbances are unexpected, the sense that one has to exert control would have its origins in signals indicating a discrepancy between predicted and actual state and in the corrections and adjustments these signals would trigger.

Typically, although not always, the more one feels one is in control the less one feels one has to exert control and vice versa. Yet, I think it would be a mistake to dispense with one of these two notions in favor of the other, for they seem to make rather different contributions to the sense of agency. Nahmias (2005) remarks that the phenomenology of effortless control is complex and somewhat ambiguous. A feeling of effortless control can sometimes give rise to a heightened sense of one's agency and sometimes involve a reduced sense of agency. One way to make sense of this ambiguity is in terms of the distinction between the feeling that one is in control and the feeling that one has to exert control. We typically experience a feeling of effortless control when we achieve a perfect match between action and goal without having to go through corrections or adjustments. So in a way our sense of agency is heightened since the performed action fully conforms to our intention. Yet, at the same time, in such actions we meet with no resistance, either internal or external, and do not experience the kind of contrast between what we want and what the world will allow that would sharpen our sense of self. In contrast, in actions where we meet with resistance and have to overcome perturbations, the actual consequences of our actions do not match our predictions perfectly and in that respect we do not feel that what we did was exactly what we wanted to do. Yet, at the same time, our awareness of the efforts we have to make to try and keep the action on track heighten our sense that we are indeed engaged in action.[16]

[16] Another variable of interest in explaining the peculiar phenomenology of effortless control may be the focus of attention of the agent. In particular, it would be worth exploring whether there is a correlation between the feelings of heightened as opposed to reduced agency for effortless actions and the inward-looking (movement of the body) or outward-looking (external goal) nature of the action.

Author's personal copy

Interestingly, executive control and feeling of mental effort are dissociable. Naccache and colleagues (Naccache et al., 2005) report the case of a woman with a left mesio-frontal cortex lesion including the anterior cingulate cortex (ACC). This patient, RMB, and a group of comparison subjects, were tested on a Stroop task, where subjects have to respond according to the ink color of a color world. In congruent trials, the ink color and the word itself refer to the same response; in incongruent trials the subject has to focus his executive attention to select the relevant information (ink color) and to inhibit the prepotent response associated with the irrelevant information (the printed color word). In normal patients, response times are slower for incongruent trials, they report feelings of subjective effort and these feelings correlate with higher skin-conductance responses (SCRs). In contrast, RMB experienced no conscious feeling of mental effort and showed no SCR, despite exhibiting normal executive control. Yet she understood the task. For instance, commenting on an incongruent trial she said: "Yes, this one was a tricky trial, with ink opposite to the word, thus it should be more difficult to me; however, I do not feel any sensation of difficulty here." (Naccache et al., 2005, p. 1323). Naccache and colleagues note that this lack of consciousness of mental effort coincides with a lack of bodily-mediated physiological responses indexing mental effort in healthy subjects. They suggest that the lesion prevented the residual activity of the right ACC, which still varied with the requirements for executive attention, from signaling internal changes in executive recruitment to the left-ventral-mesial prefrontal region known to be involved in the generation of somatic markers. On this view, the feeling of conscious effort would be a by-product of executive attention, but would not play a causal role in its deployment. Yet, Naccache and colleagues do not mean to argue that subjective feelings accompanying voluntary actions are completely epiphenomenal. RMB demonstrated a pattern of impaired behavior and SCRs in the Iowa gambling task devised by Bechara, Damasio, and Damasio (2000) suggesting that her absence of subjective feelings affected her decision-making and made her unable to progressively select the advantageous decks of cards.

In a nutshell, the account I propose of the sense of control and its contribution to the sense of agency is as follows. The degree to which one feels in control of an action is the weighted result of comparisons between desired, predicted and actual states made at three levels of action control – motor control, situational control and rational control – where comparisons at different levels may sometimes pull in different directions. The degree to which one feels one has to exert control over an action depends on the amount of adjustments and corrections one has to make to reduce the discrepancies between predictions and outcomes created by perturbations of various kinds. Control in this latter sense is normally felt as effortful, where the effort can be either mental or physical. Feeling that one is in control may heighten the sense of agency for a given action insofar as the result achieved fully conforms to the agent's intention, while feeling that one has to exert control over an action may heighten the sense that one is engaged in action, despite there being no perfect match between what is achieved and how and what was initially intended (see Fig. 7).

Author's personal copy

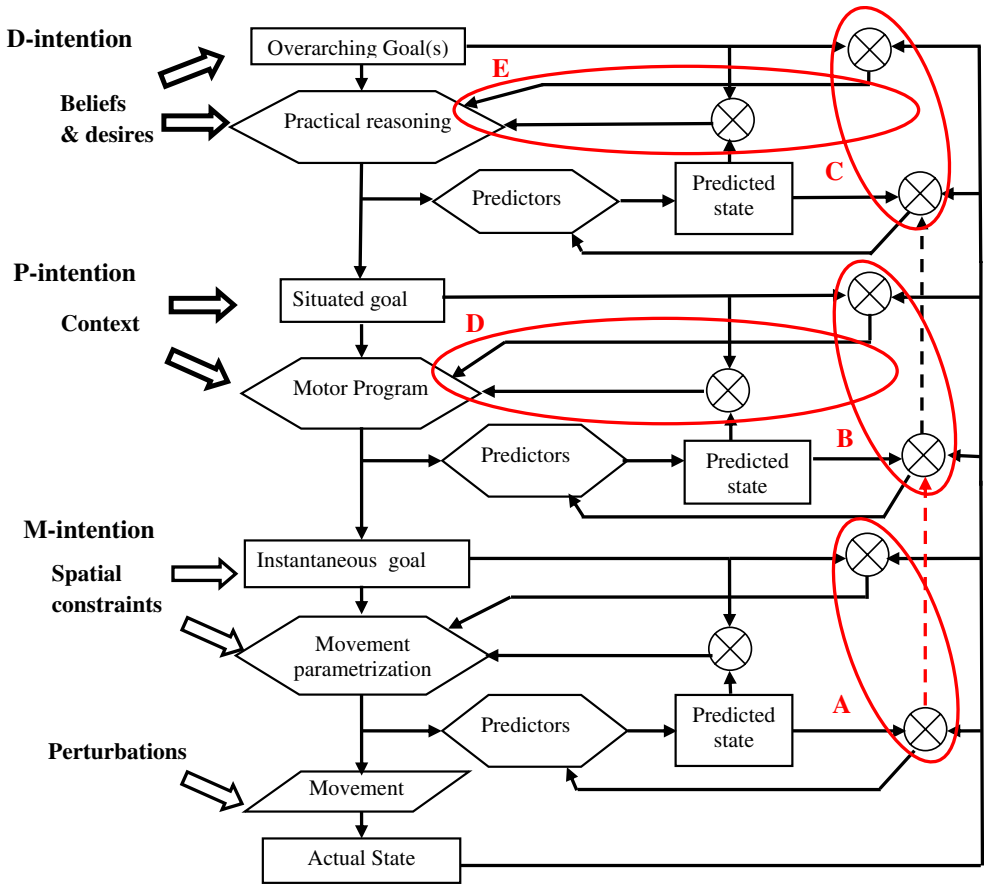


Fig. 7. The sense of control: being in control and having to exert control. The degree to which one feels in control of an action is the weighted result of comparisons between desired, predicted and actual states made at three levels of action control: motor control (A), situational control (B), and rational control (C). The degree to which one feels one has to exert control over an action and the situational (D) or rational (E) nature of this control depends on the amount of adjustments and corrections one has to make to reduce the discrepancies between predictions and outcomes created by perturbations of various kinds. Control in this latter sense is normally felt as effortful.

To recap, in this section I tried to explore the main components of the sense of agency. I argued that the sense of intentionality or intentional causation – that relies on the efferent binding of an action and its effect – is probably a necessary component of the sense of agency but is clearly not a sufficient mark of self-agency in so far as it can be present when one observes actions performed by other agents. I also argued that another form of efferent binding, between intention and movement onset, may underlie the sense of initiation for an action and make a more important contribution to the sense of authorship for an action. Finally, the sense of control – typically a compound of three more basic experiences: the sense of motor control, the sense of situational control and the sense of rational control – is another crucial component of the sense of agency for an action.

Author's personal copy

7. Conclusion

In this paper I tried to show that there are important connections between processes of action specification and control and various aspects of the phenomenology of action. I investigated the main components of our awareness of action and our sense of agency and explored their links to component representations and comparisons generated at various stages of the process of action specification. I pointed out that both awareness of action and sense of agency have multiple sources, some more important than others, and that these sources may contribute differentially according to the nature of the action – whether it is outward- or inward-directed, skilled or not, preceded by a conscious intention or not.

Although this paper investigated the possible sources of the phenomenology of agency, it said next to nothing as to the possible causal role of awareness of action or feelings of agency. If, as Wegner sometimes seems to argue (Wegner & Wheatley, 1999), the processes through which the phenomenology of agency is generated were completely separate from the processes involved in action specification and control, it would indeed be unlikely that conscious experience plays a causal role in the production and control of action. Arguing that the two sets of processes are linked keeps this option open but should not lead one to embrace without qualifications the view that our experience of doing in all its aspects systematically plays a role in the production, guidance and control of action. Rather than try and defend an all or none view, we should, as the case reported by Naccache et al. suggests, be open to the idea that some but not all of the processes of action production and control depend on conscious experiences and be ready to acknowledge the complexity of cognitive–experiential interactions.

Finally, the fact that the processes through which the phenomenology of agency is generated are closely connected to the processes involved in action specification and control suggests that our experience of doing is generally reliable but certainly not that it is infallible. Both action control and awareness of agency rely on predictions made by internal models. That reality conforms to these predictions may on occasion happen by luck rather than through our agency. In such cases, our experience of agency will be illusory. Yet the success of our predictions cannot in the general case be a matter of luck, for, given their links, this would mean not just that our sense of agency is illusory, but that agency itself is an illusion.

References

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences, 19*(1), 107–154.

Bayne, T. J., & Levy, N. (2006). The feeling of doing: Deconstructing the phenomenology of agency. In N. Sebanz & W. Prinz (Eds.), *Disorders of volition*. Cambridge, Mass: MIT Press.

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision-making and the orbito-frontal cortex. *Cognition, 50*(1–3), 7–15.

Blakemore, S.-J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Neuroscience, 2*, 561–567.

Author's personal copy

Blakemore, S.-J., Oakley, D. A., & Frith, C. D. (2003). Delusions of alien control in the normal brain. *Neuropsychologia, 41*, 1058–1067.

Blakemore, S.-J., & Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Experimental Brain Research, 153*, 239–245.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensations. *Nature Neuroscience, 1*, 635–640.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1999). Spatiotemporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience, 11*, 551–559.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Science, 6*(6), 237–242.

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

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

Bridgeman, B. (1989). Separate visual representations for perception and for visually guided behavior. In S. Ellis, M. Kayser, & A. Grunwald (Eds.), *Spatial instruments and spatial displays* (pp. 40.1–40.16). Mountain View, CA: NASA.

Buekens, F., Maesen, K., & Vanmechelen, X. (2001). Indexicaliteit en dynamische intenties. *Algemeen Nederlands Tijdschrift voor Wijsbegeerte, 93*, 165–180.

Campbell, J. (2002). *Reference and consciousness*. Oxford: Oxford University Press.

Castiello, U., Paulignan, Y., & Jeannerod, M. (1991). Temporal dissociation of motor responses and subjective awareness. A study in normal subjects. *Brain, 114*, 2639–2655.

Davidson, D. (1980). *Essays on actions and events*. Oxford: Oxford University Press.

Decety, J., Jeannerod, M., Durozard, D., & Baverel, G. (1993). Central activation of autonomic effectors during mental simulation of motor actions. *Journal of Physiology, 461*, 549–563.

Decety, J., & Michel, F. (1989). Comparative analysis of actual and mental movement times in two graphic tasks. *Brain and Cognition, 11*, 87–97.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., et al. (1994). Mapping motor representations with PET. *Nature, 371*, 600–602.

Falvey, K. (2000). Knowledge in intention. *Philosophical Studies, 99*, 21–44.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, Mass: MIT Press.

Fogassi, L., & Gallese, V. (2002). The neural correlates of action understanding in non-human primates. In M. I. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 13–35). Amsterdam: John Benjamins.

Fourneret, P., & Jeannerod, M. (1998). Limited conscious monitoring of motor performance in normal subjects. *Neurospychologia, 36*, 1133–1140.

Frankfurt, H. G. (1978). The problem of action. *American Philosophical Quarterly, 15*, 157–162.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove E. Sussex: Lawrence Erlbaum Associates.

Frith, C. D. (2005). The self in action: Lessons from delusions of control. *Consciousness and Cognition, 14*(4), 752–770.

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London B, 355*, 1771–1788.

Goldman, A. (1970). *A theory of human action*. Englewood Cliffs, NJ: Prentice-Hall.

Goodale, M. A., Pélisson, D., & Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature, 320*, 748–750.

Grèzes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation and verb generation of actions: A meta-analysis. *Human Brain Mapping, 12*, 1–19.

Grèzes, J., & Decety, J. (2002). Does visual perception afford action? Evidence from a neuroimaging study. *Neuropsychologia, 40*, 1597–1607.

Grèzes, J., Tucker, M., Armony, J., Ellis, R., & Passingham, E. E. (2003). Ofjects automatically potentiate action: An fMRI study of implicit processing. *European Journal of Neuroscience, 17*, 2735–2740.

Haffenden, A. M., & Goodale, M. A. (1998). The effect of pictorial illusion on perception and prehension. *Journal of Cognitive Neuroscience, 10*, 122–136.

Author's personal copy

Haggard, P. (2003). Conscious awareness of intention and of action. In J. Roessler & N. Eilan (Eds.), *Agency and self-awareness* (pp. 111–127). Oxford: Oxford University Press.

Haggard, P., & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition, 12*, 695–707.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382–385.

Haggard, P., & Eimer, M. (1999). On the relation between brain potentials and the awareness of voluntary movements. *Experimental Brain Research, 126*, 128–133.

Haggard, P., & Magno, E. (1999). Localizing awareness of action with Transcranial Magnetic Stimulation. *Experimental Brain Research, 127*, 102–107.

Hornsby, J. (1980). *Actions*. London: Routledge & Kegan Paul.

Humphrey, N. (1992). *A history of the mind*. New-York: Simon & Schuster.

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

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences, 17*, 187–246.

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

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

Jeannerod, M., & Pacherie, E. (2004). Agency, simulation and self-identification. *Mind & Language, 19*(2), 113–146.

Jordan, M. I., & Wolpert, D. M. (1999). Computational motor control. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences, 8*, 529–566.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activities (readiness potential): The unconscious initiation of a freely voluntary act. *Brain, 106*, 623–642.

Marcel, A. (2003). The sense of agency: Awareness and ownership of action. In J. Roessler & N. Eilan (Eds.), *Agency and self-awareness* (pp. 48–93). Oxford: Oxford University Press.

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

Mellors, C. S. (1970). First-rank symptoms of schizophrenia. *British Journal of Psychiatry, 117*, 15–23.

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

Moore, J., & Haggard, P. (in press). Awareness of action: Inference and prediction. *Consciousness and Cognition*.

Naccache, L., Dehaene, S., Cohen, L., Habert, M.-O., Guichart-Gomez, E., Galanaud, D., et al. (2005). Effortless control: Executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia, 43*, 1318–1328.

Nahmias, E. (2005). Agency, authorship and illusion. *Consciousness and Cognition, 14*, 771–785.

Pacherie, E. (1996). La schizophrénie comme déréglement de l'action – Compte rendu critique de C.D. Frith, *Neuropsychologie cognitive de la schizophrénie*. In I. Joseph, & J. Proust (Eds.) *La raison dans la place, Raisons pratiques, Vol. 7* (pp. 285–295).

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

Pacherie, E. (2001). Agency lost and found. *Philosophy, Psychology and Psychiatry, 8*(2-3), 173–177.

Pacherie, E. (2003). La dynamique des intentions. *Dialogue, XLII*, 447–480.

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., Green, M., & Bayne, T. (2006). Phenomenology and delusions: Who put the 'alien' in alien control? *Consciousness and Cognition, 15*, 566–577.

Pisella, L., Arzi, M., & Rossetti, Y. (1998). The timing of color and location processing in the motor context. *Experimental Brain Research, 121*, 270–276.

Rossetti, Y., & Revonsuo, A. (Eds.). (2000). *Beyond dissociation – interaction between dissociated implicit and explicit processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Author's personal copy

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of ownership. *Cognition, 94*, 241–255.

Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., et al. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nature Neuroscience, 7*(1), 80–84.

Slachewsky, A., Pillon, B., Fourneret, P., Pradat-Diehl, P., Jeannerod, M., & Dubois, B. (2001). Preserved adjustment but impaired awareness in a sensory-motor conflict following prefrontal lesions. *Journal of Cognitive Neuroscience, 13*, 332–340.

Spence, S. A. (2001). Alien control: From phenomenology to cognitive neurobiology. *Philosophy, Psychiatry, & Psychology, 8*(2/3), 163–172.

Spence, S. A., Brooks, D. J., Hirsh, S. R., Liddle, P. F., Meehan, J., & Grasby, P. M. (1997). A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain, 120*, 1997–2011.

Tucker, M., & Ellis, R. (1998). On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 830–846.

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D. M. (2005). Who is the controller of controlled processes? In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 19–36). Oxford: Oxford University Press.

Wegner, D. M., Fuller, V. A., & Sparrow, B. (2003). Clever hands: Uncontrolled intelligence in facilitated communication. *Journal of Personality and Social Psychology, 85*(1), 5–19.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology, 86*(6), 838–848.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation. Sources of the experience of will. *American Psychology, 54*, 480–492.

Wohlschläger, A., Engbert, K., Haggard, P., Clark, S., & Kalogeras, J. (2003). Intentionality as a constituting condition for the own self and selves. *Consciousness and Cognition, 12*(4), 708–716.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences, 1*(6), 209–216.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 358*(1431), 593–602.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience Supplement, 3*, 1212–1217.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269*, 1880–1882.