The Sense of Control and the Sense of Agency

Elisabeth Pacherie

PSYCHE

The Sense of Control and the Sense of Agency

Elisabeth Pacherie
Institut Jean Nicod
CNRS-EHESS-ENS
Département d'Etudes Cognitives
Ecole Normale Supérieure
29, rue d'Ulm
75005 Paris, France
Elisabeth.Pacherie@ens.fr

PSYCHE 13/1, April 2007

KEYWORDS: phenomenology of action; intention; action specification; motor control; sense of agency; sense of control

Abstract:. The now growing literature on the content and sources of the phenomenology of first-person agency highlights the multi-faceted character of the phenomenology of agency and makes it clear that the experience of agency includes many other experiences as components. This paper examines the possible relations between these components of our experience of acting and the processes involved in action specification and action control. After a brief discussion of our awareness of our goals and means of action, it will focus on the sense of agency for a given action, understood as the sense the agent has that he or she is the author of that action. I argue that the sense of agency can be analyzed as a compound of more basic experiences, including the experience of intentional causation, the sense of initiation and the sense of control. I further argue that the sense of control may itself be analysed into a number of more specific, partially dissociable experiences.

1. Introduction

Until quite recently, the phenomenology of action received surprisingly little attention from both action theorists and theorists of consciousness. Things are starting to change, however. In particular, improving psychological and neuroscientific methods have now made the phenomenology of agency 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 agency. At the same time, further empirical investigations aimed at probing in more detail the

phenomenology of agency and its disorders have started yielding a wealth of new data, suggesting that the current strand of will-skepticism may rest in part on too simplistic a view of the phenomenology of agency.

The now growing literature on the content and sources of the phenomenology of first-person agency highlights the multi-faceted character of the phenomenology of agency and makes it manifest that the experience of agency includes many other experiences as components. Yet, it remains unclear how these various aspects of the phenomenology of agency are linked, 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.

In this paper, my focus will be on one essential dimension of the phenomenology of doing, namely the sense of agency. The sense of agency for a given action; i.e. the sense the agent has that he or she is the author of that action, can, I shall argue, be analyzed as a compound of more basic experiences. Most prominent among these component experiences are the experience of intentional causation, the sense of initiation and the sense of control. I will further argue that the sense of control may itself be analysed into a number of more specific, partially dissociable experiences. It may therefore take different, stronger or weaker forms, depending on what, in a given instance, its sources are and on their degree of congruence.

I shall start by laying out some central assumptions regarding the relation of the phenomenology of agency to action specification and action control mechanisms that guide the approach pursued in this paper (section 2). I will then propose a preliminary regimentation of the various components of the phenomenology of agency (section 3). After a brief discussion of our awareness of our goals and means of action (section 4), I will turn to the sense of agency (section 5). I will discuss some approaches to the sense of agency one finds in the recent literature and explore the contributions the sense of intentional causation, the sense of initiation and the sense of control make to the general sense of agency, their possible relations to different aspects and stages of the processes of action specification and control, and the different forms the sense of agency may take as a result of their combined contributions.

2. Working assumptions

The approach I will pursue in order to get a better understanding of the phenomenology of agency relies on a set of assumptions that need to be made explicit. My key assumption is that the processes through which the phenomenology of agency is generated have strong connections with the processes involved in action specification and control. More specifically, the latter processes have a causal/teleological quality in the sense that representations of action goals cause general preparation, then progressive specification, then physical movement. The component representations that lead to action evolve over measurable time, and can be distinguished from each other by the time of their activation as well as their functional and content properties. Finally, these component representations are differentially accessible to consciousness, and the source of different varieties of conscious experience all linked to action.

As I have argued elsewhere (Pacherie, 2000, 2003, 2006), three main stages can be distinguished in the process of action specification, corresponding to the formation of future-directed intentions (F-intentions), present-directed intentions (P-intentions) and motor intentions (M-intentions). F-intentions are formed before the action and represent the whole action as a unit. They are usually detached from the situation of action and specify types of actions rather than tokens. Their content is therefore conceptual and descriptive. F-intentions are also, as Bratman (1987) points out, subject to distinctive normative pressures for consistency and coherence: in particular, they should be means-end coherent, consistent with the agent's beliefs and consistent with other intentions he or she may have. P-intentions serve to implement action plans inherited from F-intentions. They anchor the action plan both in time and in the situation of action and thus effect a transformation of the descriptive contents of the action plan into perceptual-actional contents constrained by the present spatial as well as non-spatial characteristics of the agent, the target of the action, and the surrounding context. The final stage in action-specification involves the transformation of the perceptual-actional contents of P-intentions into sensorimotor representations (M-intentions) through a precise specification of the spatial and temporal characteristics of the constituent elements of the selected motor program.

Another essential element of this framework is the idea that the representations formed at each of these three levels play a continuing role in the guidance and control of the ongoing action.[1] Of course, the guidance and control exercised at each level take rather different forms. As work in the area of motor control shows, for precise and smooth execution movements need to be controlled at the sensorimotor level. According to a very influential theoretical framework, motor control is achieved through the use of internal models (Jordan & Wolpert, 1999; Wolpert *et al.*, 1995; Wolpert & Ghahramani, 2000; Wolpert, 1997; Jeannerod, 1997; Frith *et al.*, 2000). The two main kinds of internal models are 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) inverse the causal flow of a system (hence their name): given a desired outcome and the current state of a system, they work out the commands that would produce the desired outcome. In motor control, inverse models capture the relationships between intended sensory consequences and the motor commands yielding those consequences. They are computational systems, which take as their inputs representations of (a) the current state of the organism (b) the current state of its environment and (c) the desired state and yield as their outputs motor commands for achieving the desired state. In contrast, the task of forward models is to predict the sensory consequences, both interoceptive and exteroceptive, of the execution of motor commands. Of special interest is the idea that the control of movement depends in a large part on the coupling of inverse and forward models through a series of comparators that compare various signals representing desired, predicted and actual states and use the result of the comparison for various kinds of regulation. For instance, discrepancies between the predicted and actual consequences of the execution of a movement can be used to instantly adjust the ongoing movement. These processes of online motor control are very fast. The sensorimotor format of the representations they exploit and their temporal properties make it doubtful whether their contents could in principle be accessible to consciousness.

In contrast, the higher-level guidance and control functions exercised at the level of P-intentions and F-intentions typically take a conscious form. They are involved in the rational and the situational control of action, supporting control processes responsible for keeping track of the way the agent accomplishes her action and adjust it so as to maximize her chances of success (tracking control) and to minimize undesirable side-effects (collateral control).[2] Here, the main difference between P-intentions and F-intentions is that the former exercize tracking and collateral control of the action with regard to the situation as currently perceived, whereas the latter are concerned with the respect of more global consistency and coherence constraints.

Thinking of the control of action in terms of internal models has proven very fruitful. Although the main in-depth application of this idea has been to fine-grained aspects of motor control, corresponding to the level of M-intentions, there is no good reason why the idea of internal models shouldn't be used in thinking about more global aspects of action specification. Presumably, the deliberative processes at work at the level of F-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 F-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. Yet, the general idea that internal models divide into inverse models which compute the means towards a given goal and forward models which compute the consequences of implementing these means retains its validity at the level of F-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 differ from both F-level and P-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 F-intentions. On the other hand, 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.[3]

I therefore suggest that the information-processing model of action control in terms of internal models be explicitly combined with the threefold distinction among kinds of intentions I tried to motivate, thus yielding a richer theoretical framework for thinking about action. Figure 1 provides a schematic representation of the view of action specification and control that results from this combination.

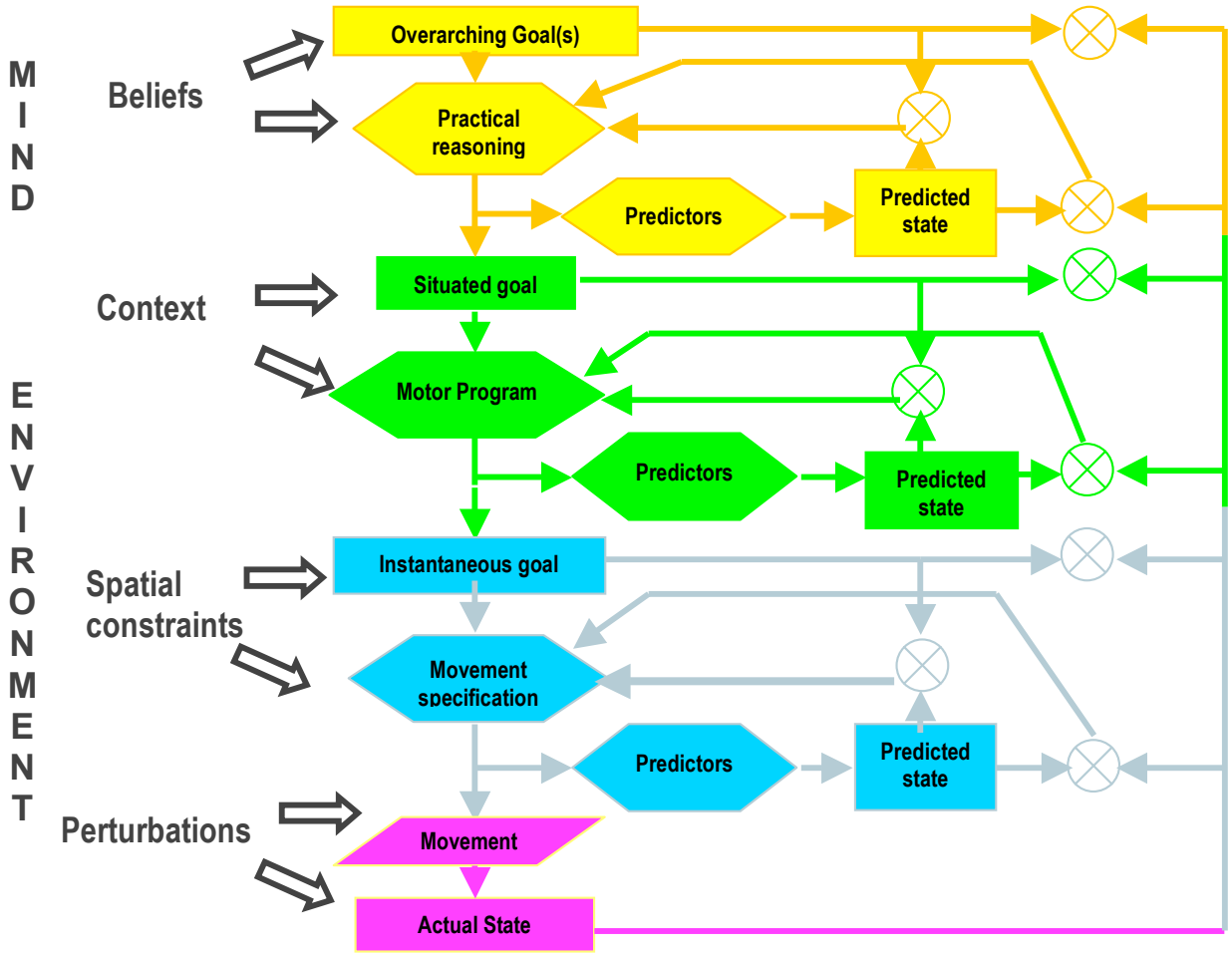

Figure 1: A hierarchical model of action specification. The top level, in yellow, corresponds to F-intentions, the middle level in green to P-intentions, the level in blue to M-intentions, and the level in pink to actual execution.

What I now want to explore is the idea that some at least of these information-processing events may have phenomenal counterparts and that it may be possible to identify links between conscious experiences during voluntary action and action-specification processes by considering their respective contents and temporal properties. Although, I'll briefly discuss other aspects of the phenomenology of agency, my main focus will be on the experience of control.

3. A preliminary regimentation

Both philosophical and empirical investigations highlight the fact that the phenomenology of agency has 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. This profusion raises several questions: how are these various aspects of the phenomenology of agency related? To what extent are they are dissociable? Are some more basic than others? Where does their content come from? 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 first 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, don't. 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.

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.[4] 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 a sense of agency that is both occurrent and immersed.

Yet, even this minimal sense of agency 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 agency, 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're 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 relate to one another and how they could relate to component representations built at various stages of the process of action-specification.

4. Awareness of action

4.1. 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 of what we are doing. According to the model of action specification described earlier, the goal of an action can be specified at three levels. When the action is preceded by deliberation and we have an F-intention, we can be conscious of an intended goal prior to any situational anchoring and independently of the actual performance of the action. To take a very simple example, I might form the intention to get a drink of water and be aware of that goal. At this level, the goal is represented in a rather abstract and coarse-grained way, as a goal of a certain type or belonging to certain semantic category, for instance drinking as opposed to eating, or drinking water as opposed to soda or milk.

That we can be aware of an intended goal as represented in the content of a F-intention doesn't mean that the experience of purposefulness always takes that form. For one thing, actions are often performed without prior deliberation and without being preceded by F-intentions. Another common occurrence is the phenomenon of lost intentions. I may find myself in the kitchen wondering why on earth I went there. Yet, my inability to remember that I intended to get a glass of water is not enough to completely undermine a sense of purposiveness for my action. Presumably, the experience one has in such a case is rather different from the experience of a sleepwalker suddenly awakened who finds himself in his kitchen. It may be that our familiarity with such slips of intentions leads us to infer the existence of an initial purpose and that this in turn helps us retain a sense of purposefulness for the action.

It may also be that that our sense of purposiveness in the case of actions performed without being preceded F-intentions or in the case of slips of intentions derives in part from our awareness of the goal as it is represented at the level of P-intentions. At this level the goal is represented in a more specific way, not just semantically but indexically, not just for instance as "to get a drink of water" but as "to reach for this glass of water". We are usually aware of our situated goal both immediately prior to action initiation and while the action unfolds. Thus, although I may have forgotten why I wanted to go to the kitchen, I may still be aware that I wanted to go to the kitchen. In other words, one can lose sight of one's overarching goal, while remaining aware of one's immediate situated goal, which in the example considered here, happens to be a subgoal to the overarching goal of getting a drink. I'll say more below on the relationship between goals at the level of F-intentions and goals at the level of P-intentions. The point for now

is simply that the sense of purposiveness for an action may be linked to awareness of the goal at the level of P-intentions rather than, or in addition to, awareness of the goal at the level of F-intentions.

We are often but not always aware of the situated goal of our actions. To take a well-rehearsed example, in highly routine or automatic actions, such as driving a car, one may more or less automatically respond to stimuli in the environment without being aware (or while being only marginally aware) of what we are doing. Some of our goals may be explicitly conscious but not salient, say taking a left turn after the gas station when following a well-known route; some may not be explicitly conscious at all, such as slowing down before a curve or shifting gears. Yet, they can be principle become explicitly conscious. This happens, for instance, in case of action failure. If we apply the brakes because the truck in front of us is slowing down or turn the wheel to take a curve and the car doesn't respond, we may suddenly become aware of what we were trying to do.

It is unclear whether we can be aware of our instantaneous motor goals, as they are implemented in a sensori-motor format. F-intentions, P-intentions and M-intentions employ different representational formats and the representations at these levels have different temporal properties. It is therefore possible that the representational format of M-intentions is incommensurable with the representational format or formats of our phenomenology or 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 optic ataxia. For instance, the patient DF, studied by Milner an 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 visuomotor 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 visuomotor abilities suggests that visuomotor representations are need not be derived from conscious visual perceptions but can be built independently. Conversely, it also suggests that conscious visual representations cannot be derived from intact sensori-motor representations.

Yet, as Wakefield and Dreyfus point out: ‘Although at certain times during an action we may not know what we are doing, we do always seem to know during an action that we are acting, at least in the sense that we experience ourselves as acting rather than as being passively moved about’ (1991: 268). It seems that even when awareness of action reduces to the sense that we are acting and does not include an awareness of what we are doing, a minimal sense of purposiveness is retained despite the specific goal of the action not being conscious itself. I suggest that M-intentions may still be responsible both for this basic aspect of action phenomenology, that-experience, and for the minimal sense of purposiveness that tinges it.

As we have seen, motor control involves mechanisms of action anticipation and correction. Although, these mechanisms largely operate at the subpersonal level, in the sense that the representations they process are typically unavailable to consciousness, they may nevertheless underlie the experience of acting in its most basic form. In other words, our awareness that we are acting, the sense of bodily implication we experience may result from the detection by the comparison mechanisms used in motor control of a

coherent sensory-motor flow. It is important to note that on this view, the basic experience that one is acting need not involve conscious access to the *contents* of the sensorimotor representations used for the control of the ongoing action but only to the result of the comparisons between intended, predicted and actual states. That is why, as Wakefield and Dreyfus remark, we may experience ourselves as acting without knowing what it is exactly we are doing. That is also why the representational content of the experience of acting may appear so thin. This is especially so when one is engaged in 'minimal' actions, actions that are performed routinely, automatically, impulsively or unthinkingly. These actions unfold with little or no conscious control by P-intentions. Their phenomenology may therefore involve nothing more than the faint phenomenal echo arising from coherent sensory-motor flow.

In a nutshell, I suggest that the sense of purposiveness that forms part of the awareness of actions has three main sources: awareness of the goal of the action as specified at the level of F-intention, awareness of the situated goal as specified at the level of P-intentions, basic sense of doing that arises from a comparison of intended, predicted and actual states at the level of M-intentions. The sense of purposiveness accompanying an action can be thinner or thicker depending on how many of these sources contribute in a given case and how much they do. For instance, although the goal of an action as represented at the level of F- and P-intentions is in principle accessible to consciousness, it may not be explicitly conscious at a given time if attention is focused elsewhere, the agent is distracted, or the action is routine and automatic.

4. 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 F-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 notion of basic actions. As argued by Hornsby (1980), one should at least distinguish between a causal notion of basicness and an intentional (or teleological) notion of basicness. Hornsby offers the following definition of causal basicness:

> A description *d* of a particular action *a* is a more basic$_C$ description than another description *d′* if the effect that is introduced by <*d, a*> causes the effect that is introduced by <*d′, a*>. (1980: 71).

While her notion of teleological or intentional basicness runs as follows:

> The kinds of actions in an agent's repertoire that are basic for him are those which he knows how to do, and knows how to do otherwise than on the basis of knowing how they are done by him (1980: 84)

Three important features of this latter definition are worth stressing. First, the idea of an intentionally basic action is related to that of know-how. Second, knowing how to do something should not be taken to imply that one possesses theoretical knowledge of how it is done. This know-how should be considered as immediate, that is non-mediated by a theoretical knowledge of how it is done. In other words, it is practical and not theoretical know-how. Third, Hornsby's notion of intentionally basic action is relativized to individual agents. Thus, for example, executing a trill may be a basic action for a professional pianist, but not for a novice piano player.

One important reason for drawing a distinction between these two notions of basic actions is that causal basicness and intentional basicness do not always go hand in hand. For instance, if a man raises his right arm, 'contracting such and such muscles' will be a description of his action that is causally more basic than 'raising his arm', but, unless perhaps the man is a yogi or an expert body-builder, contracting his muscles is not something on which he will have direct voluntary control and thus the description of his action as 'raising his arm' will be intentionally more basic than the description 'contracting such and such muscles'.

My reasons for drawing attention to this distinction 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 F-intentions are typically the subgoals in achieving the goal and these subgoals are represented in an abstract, semantic way. When the F-intention is for a familiar action, say, going to work, we frequently don't 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 F-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 F-intentions become immediate situated goals while means correspond to ways of implementing these goals, the selected motor programs and the types of movements or movement sequences they define.

So to what extent are we aware of our bodily movements? Awareness of movements appear 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 basic for them and have 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 for 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 and 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 couldn't 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, such that the

visible trajectory no longer corresponded to that of the hand, and the bias was increased from trial to trial. In order to reach the target, the hand-held stylus had to be moved in a direction opposite to the bias. In other words, although the line on the computer screen appeared to be directed to the target location, the hand movement was directed in a different direction. 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.

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 the direction of their hand. 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.[5] The general idea suggested by this result is that when biases remain small enough the visuomotor system is able to appropriately use information for producing accurate corrections to reach a target, but that this information is not accessed consciously. However, when the bias exceeded a mean value of about 14 degrees, 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 experienced their movements either 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 doesn't 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'. 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.[6]

One reason for preferring this alternative 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 can not 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 the reach it. Subjects were asked (a) to draw with their 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 to 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. A further 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 *et al.* (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, 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.

5. Sense of agency

At first blush, it may appear strange to consider separately the question of awareness of action and of sense of agency for our actions, where the sense of agency is defined as the sense that one is the author of that action. Indeed, 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 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 actions, 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: 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. Before proposing my own account, I discuss three approaches to the sense of agency one finds in the recent literature.

5.1. Proprioceptive awareness

According to the first approach, the primary source of the sense of agency is the sensory modality carrying the information about the action. Exteroceptive perceptual modalities, such as vision or audition, carry information about both our actions and the actions of others. In contrast, information gained through proprioception is guaranteed to be about our actions. Thus, awareness of the modality through which we are informed of an action would give us the identity of the agent. If by proprioception, it must be oneself. They are problems though with this account of the sense of agency in terms of proprioceptive awareness. First, as we already saw, it seems that we have little proprioceptive awareness of actions where the primary aim is to achieve an external goal as opposed to actions where the primary aim is to make a movement. Thus, this approach in terms of proprioceptive awareness may perhaps account for our sense of agency for the latter kind of actions, but it is doubtful it would generalize to all actions. Of course, in actions directed at achieving an external goal, we may still have some proprioceptive awareness that we are moving, but this awareness would presumably be too coarse-grained to be matched to a specific action. Suppose, for instance, I am visually attending to someone else's action while acting myself, it is unclear which of these two actions my coarse-grained proprioceptive awareness that I am moving should give me a sense of agency for. Second, schizophrenic agents experiencing delusions of control do not deny, say, that their arm that is moving, what they deny is being the agent of that action. In other words, they do not deny a sense of ownership of the movement, but they deny a sense of authorship of the action. This suggests that proprioceptive awareness may underlie the sense of ownership for movements, but may not be sufficient for a sense of authorship of

actions, in other words for a sense of agency.[7] Proprioception may not even be necessary for the sense of agency. As Marcel (2003) points out, this is shown by the case of Ian Waterman, reported by Cole (Cole, 1993; Cole & Paillard, 1995). Through a peripheral neuropathy, IW is completely deafferented below the neck and hence deprived of all proprioceptive experience beneath the neck. When he does something, he cannot tell without visual feedback the disposition of his limbs and body, yet he knows that he has acted.

Instead of seeing proprioceptive awareness as underlying the sense of agency for one's action, one could be tempted to take the opposite view and to argue, as Frith (2005) does, that *lack* of proprioceptive awareness could be a possible indicator that I am performing a voluntary act. Numerous studies have shown that the perceptual consequences of self-generated actions are attenuated (Blakemore *et al.*, 1998; 1999, 2000, 2002; Claxton, 1975; Collins *et al.*, 1998). Blakemore and colleagues suggest that proprioceptive feedback is attenuated during voluntary movements through forward modeling. The idea is that we can predict the feedback we will receive on the basis of the motor commands we are preparing to issue and use these predictions to filter incoming sensory information and thus attenuate self-produced sensory stimulation. One important line of evidence in favor of this view are their extensive studies of the phenomenon of self-tickling (Blakemore *et al.*, 1998, 1999, 2000, 2002) that show that tickling sensations are attenuated both phenomenologically and physiologically and that the amount of attenuation is proportional to the spatial and temporal congruence of the predicted and actual feedback. However, a recent study by Tsakiris and Haggard (2003) shows that sensory attenuation is present whenever an effect is self-generated even when the motor system cannot accurately predict the details of the sensory consequences of the performed action. Tsakiris and Haggard (2003, 2005) therefore suggest that the sensory attenuation of self-generated effects could have two complementary sources: The generation of efference would be sufficient to produce an attenuation of subsequent sensory effects by a fixed amount and then detailed predictions of sensory consequences made by forward models, when available, would yield a further attenuation of sensory effects in a proportional manner.

Lack of proprioceptive experience may, as Frith suggests, contribute to the sense of agency for an action. Yet, it can't be the main index that an agent is engaged in voluntary action, for presumably we also lack proprioceptive experience when we are neither actively moving nor being passively moved.

5.2. Awareness of intentions

A second possible approach to the sense of agency is in terms of awareness of intentions. Both Humphrey (1992) and Wegner (2002) argue that the sense of agency[8] is inferred from the existence of a match between a prior intention and an observed action. Frith's early account of alien control in schizophrenia (Frith, 1987, 1992) went along similar lines. The main components of his account were a distinction between two kinds of intentions — stimulus intentions (i.e., unconscious intentions automatically triggered by a stimulus and willed intentions (i.e., conscious intentions based on internal plans and goals) —, together with a distinction between two levels of monitoring. At the lower level, action-monitoring involved using efference-copying mechanisms to distinguish

between changes due to our actions and changes due to external factors. At the higher level, intention-monitoring involved a comparison of willed intentions with executed actions. His account of alien control posited the existence of a deficit in intention-monitoring resulting in the *loss of awareness of 'willed' intentions* to act. The main difference between Frith's account and Humphrey's and Wegner's lies in the fact that for Frith the comparison between intention and action is done subpersonally with only its result being consciously available in the form of a sense of agency for the action.

There are two main problems with this approach. 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. As Frith himself later acknowledged, the idea that experiences of alien control arise through a lack of awareness of intended actions "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: 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.

5.3. Intentional binding

On the approach just discussed the sense of agency arises primarily when there is match between a prior intention and an observed action. Yet, they are other types of matches that may be worth considering. One such match is between an action and its consequences. Haggard and colleagues (Haggard & al., 2002; Haggard & Clark, 2003) 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 points out that this phenomenon, which he calls intentional binding, depends critically on the intention to produce the effect. When similar movements and auditory effects occur involuntarily rather than by the subject's intention, the binding effect is reversed and cause and effect are perceived as further apart in time than they actually are.

Haggard suggests that intentional binding is best explained in terms of predictive mechanisms of action control: it depends on efferent signals since it does not occur with passive movements and it causes anticipatory awareness of action effects, a shift that suggests prediction. On this predictive account, the conscious experience of action would 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) tested their predictive account by using TMS to insert occasional involuntary movements of the right finger at a time when the subject was intended 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 intention followed by the appropriate effect is not sufficient for intentional binding and appear incompatible with a reconstructive account of the sense of agency of the type favored by Wegner, where the existence of a match between intention and action is enough for the agent to retrospectively infer that he was the source of the action.

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. Evidence that it may not be so comes from further experimental work using the same paradigm. 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. Subjects had slightly delayed awareness of the onset of their own actions and of the experimenter's action, showing in both cases a binding effect, but an anticipatory awareness of the machine's action.

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.[9] If, however, intentional binding is not linked to a particular person, it cannot be the basis of the sense of authorship for an action. Intentional binding of action and effect would seem to be associated with the agent-neutral experience of intentional causation, rather than with the experience of authorship per se. 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.[10] This dissociation of the two experiences may indeed 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, may still be a necessary component of that experience.

In pointing out the limitations of these three approaches just discussed I did not mean to deny that proprioceptive awareness, awareness of a match between a prior intention and an action, or intentional binding can contribute to the sense of agency one normally feels for one's actions. Rather, I want to suggest that some further processes involved in the preparation and control of action may play a more crucial role. This is the approach I will now pursue. The two contributors I have in mind are sense of initiation and sense of control.

5.4. Sense of initiation

According to Frith, Blakemore and Wolpert's revised account of delusions of alien control (Frith *et al.* 2000; Blakemore *et al.* 2002, 2003; Frith, 2005), 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).[11] According to this revised 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 *et al.*, 1983; Libet, 1985).

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 they also showed that both types of judgments were unrelated to the general readiness potential but covaried with the lateralized readiness potential.[12] 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 investigating altered awareness of voluntary action after damage to the parietal cortex, 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. This is consistent with independent evidence that the parietal cortex is important in activating and maintaining internal models used to predict the future outcome of a given action (Sirigu *et al.*, 1996; Desmurget & Grafton, 2000).

As Haggard (2005) suggests, the subjective experience of conscious intention often contains two components: a sense of urge, of being about to move, and a reference forward to the goal object or event. One might speculate that awareness of intention depends on activity in the SMA for its first aspect and on parietal processing for the second, while awareness of movement would not depend on parietal activity but have its source in premotor processing in the SMA. As Haggard (2003) points out, the co-existence of awareness of intention and awareness of movement onset within a single narrow window of premotor 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 moving, 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 5.3. 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 action initiation. One should further note that this proposal differs in two important ways from the proposals of Wegner, Frith and Humphrey discussed earlier. First, the representations that are bound are not representations of a prior conscious intention (F-intention) and of an observed action, but of an immediate intention (P-intention) and of its internally predicted consequences. Second, this binding is not a post-hoc reconstruction but a preconstruction.

5.5. Sense of control

Although the sense of initiation may be a crucial component of the sense of agency for an action, it doesn't seem to offer the guarantee that the whole 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. Some experimental data provide further evidence for this possibility. Thus, in an experiment where they had to move a cursor on a computer screen, some schizophrenic patients had a normal sense of agency while the cursor was visible but experienced a strong sense of alien control when visual feedback was withdrawn, the cursor in effect moving behind an opaque obstacle on the screen (Jouvent, personal communication).

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 2, 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 F-

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 control at the level of F-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 4.2 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 sensorimotor signals used for motor control could contribute something to the experience of control. This contribution would, however, be limited in two ways: first, sensorimotor 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 F-intention, it would depend on whether our action plan is carried out successfully, in other words on the conceptual match between the predicted and actual consequences of its successive steps. At both levels, the better the match the more one feels in control. Yet, it is important to note that these two forms of the sense that one is in control—conceptual and rational at the level of F-intentions, perceptual and situational at the level of P-intentions—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.

The sense that one is in control is therefore 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 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 that when one doesn't feel in full motor control one is simply aware that something is wrong, whereas when one's

doesn't 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 look, 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 don't 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.[13]

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 where the ink color and the color word don't match, 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 heightened 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: 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 *et al.* (2000) suggesting that her absence of subjective feelings affected her decision-making and made her unable to progressively select the advantageous decks of cards.

To recap, 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 predicted and actual states made at three levels of action control—motor control, situational control and rational control—where comparisons at different levels may sometime 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.

One may object to this account that it takes things from the wrong end.[14] It claims that the sense of control for an action depends on comparisons between predicted states and actual states at various levels of action specification. This is in effect a kind of bottom-up account. One may want to claim instead that our beliefs about what we can or cannot control as well as information about other possible explanations of how some effects are produced may lead us to infer that we control an action or that we don't. According to this top-down alternative to the account defended here, rather than the sense of control stemming from comparisons between predicted states and actual states, our prior beliefs about control would influence the comparisons we make.

One line of evidence that may be thought to favor this alternative view comes from work in developmental psychology. Schultz *et al.* (1980) used a reflex hammer to

induce a knee-jerk reflex in children. The children were then asked whether they had meant to move their leg. Three-year-olds said that they had meant to move their leg but five-year-olds denied it correctly. Lang and Perner (2002) reproduced these results and also showed that performance on the knee-jerk reflex task strongly correlates with performance on the false-belief task and on two executive function tasks (card sorting task and hand game) where one has to inhibit interfering action tendencies.

The objection to the account I defend in this paper goes something like that. The question "Did you mean to move your leg?" translates as "Did you control the movement of your leg?" On my account, to answer this question, one should consult one's phenomenology so to speak. If the subject experienced a sense of control for the movement, she should say that she meant it; otherwise, she should say that she didn't mean it. If this is how things are, the account has no ready explanation why three-year-olds are not able to answer correctly and five-year old are, for presumably both groups of children enjoy the same phenomenology. In contrast, on the alternative top-down account, the performance of the children in the knee-jerk reflex task can be explained in terms of their theoretical beliefs about movements. The older children have acquired, while the younger children still lack, beliefs about the difference between reflex movements and other movements. In particular, the older children have acquired the belief that reflex movements are movements they can't control and are thus in a position to infer that they did not mean to move their leg, while the younger children haven't yet acquired the beliefs that would support this inference.

There is, however, a third, altogether more plausible, explanation of the performance of younger children on that task. Namely, the reason the younger children fail the task is that they have yet to develop a sufficient mastery of the concepts of intention and intentional action. Indeed, there is clear evidence, summarized in Astington (2001), that children do no really understand the distinction between desire and intention until they are four or five years old. For instance, younger children will tend to say that an accidental action was intended if its outcome satisfies a desire they have. They have an appreciation of desires and their relation to actions but they still lack an appreciation of the role of intentions as causal mediators between desires and actions and their outcomes.

One may want to argue that an appreciation of the causal role of intentions is also what is needed to distinguish between ownership and authorship of movements. Children who lack this appreciation would tend to treat the questions "Did your leg move?", "Did you move your leg?", and "Did you mean to move your leg?" as synonymous. On this view, passing the knee-jerk reflex task does not require the acquisition of theoretical beliefs about movements and their control, but a more developed understanding of intentions and their role. It may be as Astington (2001) and Lang and Perner (2002) would argue, and as the correlation between performance on the knee-jerk reflex task and on the false belief task suggests, that such an understanding depends on the development of metarepresentational abilities. It may also be, as Russell (1996) and I (Pacherie, 1997) have argued, and as the correlation between performance on the knee-jerk reflex task and on executive function tasks suggests, that such an understanding depends on the development of executive control. The idea is that for children to develop an understanding of the specific causal role of intentions, their intentions should indeed play this role, something that depends on the development of executive control.

The important point here is that, on either view, successful performance on the knee-jerk reflex task does not depend on the acquisition of beliefs about movements and their control but on the development of an understanding of the role of intentions. It is only once this understanding is secured, that children can translate the question they are asked as "Did you have the intention to move your leg and did this intention cause and control the movement of your leg?" and can start looking for experiential cues that will help them answer it. Both views are therefore compatible with a bottom-up account of sense of control.

I don't think therefore that the existing developmental evidence supports the view that the sense of control is primarily the result of a theoretical inference that would then influence comparisons. There are, however, two things to be said in favor of top-down effects. First, the phenomenology of action in general and the experience of control in particular are often thin and elusive, with minimal, ephemeral conscious content. Top-down influences may therefore enhance otherwise elusive experiences through a process of attentional amplification. Thus, in younger children the constituent elements of the experience of control may already be present and yet go unnoticed, while older children with a better grasp of the concept of intention may attend to their experience, thus increasing its salience. Note though that younger children also have limited executive function abilities, hence a limited control over their actions and presumably a reduced experience of control. There is also a second way in which top-down effects may influence the sense of control. As both philosophers and psychologists have argued (Brand, 1984; Bratman, 1987; Davidson, 1980; Moses, 2001), besides a motivational and a causal aspect, intentions also have an epistemic aspect. An intention to perform an action normally entails a belief that one will perform the action. There are divergences as to the exact form such beliefs should take and how strong they should be, but there is at least a consensus that I cannot intend what I positively believe that I can't do. Thus it seems that beliefs about what we can and cannot control might influence not just our intention-formation processes but also our interpretation of what we actually do. In the absence of unambiguous experiences of motor and situational control, whether or not I experience control for an action may well depend on my prior beliefs about control. This, however, doesn't make these beliefs the primary determinants of my sense of control for an action; rather they would provide auxiliary sources of information used for disambiguation.

6. Conclusion

What I tried to do in the paper is show that there are important connections between processes of action-specification and control and various aspects of the phenomenology of action. My main focus was on the sense of agency and I tried to explore its main components. 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 appears to be agent-neutral. I also argued that another form of efferent binding, between intention and movement onset, may underlie the sense of initiation for an action and play a more important contribution to the sense of authorship or an action. Finally, although the sense of initiation may be a crucial component of the sense of agency for an action, it doesn't seem to offer the guarantee that the whole action will be owned by the subject. A

sense of control seems to be also required. I tried to show that the global sense of control for a given action is a compound of three more basic experiences: the sense of motor control, the sense of situational control and the sense of rational control, each resulting from comparisons between predicted states and actual states at a given level of action specification and control. These more basic experiences are partly dissociable and may contribute differentially according to the nature of the action—for instance whether it is skilled or not, preceded by a conscious intention or not. One of the other approaches to the sense of agency I briefly discussed emphasizes the existence of a match between a prior conscious intention and an observed action. In my view, this match is but one of the elements contributing to the sense of control and not the most crucial one at that. Indeed, I concur with Haggard's view (Haggard, 2005) that the sense of agency should be seen as mostly a preconstruction rather than a post-hoc reconstruction

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 shouldn't 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 processes of action production and control depend on conscious experiences and be ready to acknowledge the complexity of cognitive-experiential interactions.

References

Astington, J. W. (2001). The paradox of intention: assessing children's metarepresentational understanding. In Malle, B. F., Moses L. J., and Baldwin D. A. (Eds.), *Intentions and Intentionality - Foundations of Social Cognition*, Cambridge, MA: MIT Press, pp. 85-103.

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. 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. (2000). Why can't you tickle yourself? *NeuroReport*, 11, R11-R16.

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

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

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

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

Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Cambridge University Press.

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

Claxton, G. (1975). Why can't we tickle ourselves? *Perceptual and Motor Skills*, 41: 335-338.

Cole, J. & Paillard, J. (1995). *Living without touch and peripheral information about body position and movement.* In J. Bermudez, A. Marcel & N. Eilan (Eds.), The Body and the Self. Cambridge, Mass.: MIT Press.

Cole, J. (1993). *Pride and a Daily Marathon*. Cambridge, Mass.: MIT Press.

Collins, D. A., Cameron, T., Gillard, D. M., & Prochazka, A. (1998). Muscular sense is attenuated when humans move. *Journal of physiology*, 508: 635-643.

Davidson, D. (1980). *Essays on Actions and Events*. Oxford University Press, Oxford.

Desmurget, M. & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Science*, 4, 423.

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. *Neuropsychologia*, 36: 1133-1140.

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological Medicine*, 17: 631-48.

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.

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.

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

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Science*, 9, 6: 290-295.

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

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.

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

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

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

Jeannerod, M. & Pacherie, E. (2004). Agency, Simulation and Self-identification. *Mind & Language*, 19, 2: 113-146.

Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Oxford: Blackwell.

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

Lang, B. & Perner, J. (2002). Understanding of intention and false belief and the development of self-control. British Journal of Developmental Psychology, 20: 67-76.

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*. Oxford: Oxford University Press, pp. 48-93.

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

Milner, A. D., and Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Moses, L. J. (2001). Some thoughts on ascribing complex intentional concepts to young children. In Malle, B. F., Moses L. J., and Baldwin D. A. (Eds.), *Intentions and Intentionality - Foundations of Social Cognition*, Cambridge, MA: MIT Press, pp. 69-83.

Naccache, L., Dehaene, S., Cohen, L., Habert, M.-O., Guichart-Gomez, E., Galanaud, D. & Willer, J.-C. (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*, 7: 285-295.

Pacherie, E. (1997). Motor images, self-consciousness and autism. In J. Russell (Ed.), *Autism as an Executive Disorder*. Oxford, Oxford University Press, pp. 215-255.

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

Pacherie, E. (2001). Agency lost and found. *Philosophy, Psychiatry and Psychology*, 8, 2-3: 173-176.

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*. Cambridge, MA: MIT Press, pp. 145-167

Pacherie, E., Green, M. & Bayne, T. (2006). Phenomenology and delusions: Who put the 'alien' in alien control?. *Consciousness and Cognition*, 15: 566-77.

Russell, J. (1996). *Agency: Its role in mental development*. Hove: Erlbaum (UK)/Taylor & Francis.

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

Searle, J. (2004). *Rationality in Action*. Cambridge, MA: MIT Press.

Shultz, T. R., Wells, D. & Sarda, M. (1980). The development of the ability to distinguish intended actions from mistakes, reflexes and passive movements. *British Journal of Social and Clinical Psychology*, 19: 301-310.

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

Sirigu, A., Duhamel, J.-R., Cohen, L., Pillon, B., Dubois, B., Agid, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science*. 273(5281): 1564-1568.

Slachewsky, A., Pillon, B., Fourneret, P. Pradat-Diehl, 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, Number 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.

Tsakiris, M. & Haggard, P. (2003). Awareness of somatic events following a voluntary action. *Experimental Brain Research*, 149, 439-446

Tsakiris, M. & Haggard, P. (2005). Experimenting with the acting self. *Cognitive Neuropsychology*, 22, 3/4, 387-407.

Velleman, D. J. (2006). *From Self to Self*. New-York: Cambridge University Press.

Wakefield, J. & Dreyfus, H. (1991). Intentionality and the phenomenology of action. In E. Lepore & R. Van Gulick (Eds.), *John Searle and his critics*. Cambridge, MA: Blackwell, pp. 259-270.

Wegner, D. M. & Wheatley, T. (1999). Apparent mental causation. Sources of the experience of will. *American Psychologist*, 54: 480-92.

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

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. & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience Supplement*, 3: 1212-1217.

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

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

Notes

[1] Although my F-intentions are somewhat similar to Searle's prior intentions (Searle, 1983), one crucial aspect in which they differ is that I do not conceive of F-intentions as ending with the onset of the corresponding body movements. Note though that Searle (2004) withdrew his earlier claim that the role of prior intentions is over once movement starts.

[2] I borrow this useful distinction between two forms of rational control, tracking and collateral control, from Buekens, Maesen and Vanmechelen (2001).

[3] 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. For a discussion, see Pacherie *et al.*, in press.

[4] For an approach of long-term agency along these lines, see for instance Velleman (2006)

[5] The mode of response used in this experiment (indicating a line on a chart) may be thought to have induced a bias and to have led 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).

[6] I used to think, together with Jeannerod, that failure of action could result in an awareness of motor representations normally outside the agent's subjective experience. We speculated that the main reason these representations are not normally consciously accessible is that that they are too short-lived. In the case of a successfully executed motor act, the content of the motor representation would not reach consciousness because it would be cancelled as soon as the corresponding movements were executed (perhaps by the incoming signals generated by the execution itself). In contrast, in cases where execution is blocked or error corrections fail the representation would be protected from cancellation and would become accessible to conscious processing. Yet, whether a representation is accessible to consciousness or not may not just be a matter of its temporal characteristics but also of its representational format. It is therefore doubtful whether sensori-motor representations as such can in principle be accessible to consciousness.

[7] It would be interesting to know whether schizophrenic patients experience delusions of control primarily for actions directed at an external goal or primarily for actions aimed at making a movement. I am aware of no data bearing on that issue, but such data would presumably give us some indication on the contribution, important or not, proprioception makes to the sense of agency.

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

[9] Indeed, there is increasing evidence that common brain circuits are used both to control one's actions and to represent observed actions performed by others. This evidence ranges from single-cell recordings studies in monkeys, where mirror neurons were discovered that fire both during goal-directed action and observation of actions performed by another individual (see Fogassi & Gallese, 2002 for a review) to functional neuroimaging experiments in humans (see Blakemore & Decety, 2001 and Grèzes & Decety, 2001, for reviews).

[10] 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.

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

[12] Intention judgments correspond to what Libet, and Haggard following Libet, 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.

[13] 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.

[14] I am grateful to Joshua Knobe for drawing my attention to this objection.