

For *Contemporary Debates in the Philosophy of Mind* (Blackwell, 2007), ed. J. Cohen and B. McLaughlin

Mental Action and Self-Awareness (I)¹

Christopher Peacocke

This paper is built around a single, simple idea. It is widely agreed that there is a distinctive kind of awareness each of us has of his own bodily actions. This action-awareness is different from any perceptual awareness a subject may have of his own actions; it can exist in the absence of such perceptual awareness. The single, simple idea around which this paper is built is that the distinctive awareness that subjects have of their own mental actions is a form of action-awareness. Subjects' awareness of their own mental actions is a species of the same genus that also includes the distinctive awareness of bodily actions. More specifically, I claim:

(1) Much conscious thought consists of mental actions.

(2) A thinker's awareness of those of his mental events that are mental actions is a species of action-awareness. This I call "The Principal Hypothesis".

(3) The Principal Hypothesis can provide a clarification and explanation of a range of features and phenomena present in conscious thought.

(4) The Principal Hypothesis is a resource that can be used in addressing various classical philosophical issues about the mental, self-knowledge, and the first person.

¹ Earlier versions of this paper were presented in 2004 to the Language and Mind Seminar at NYU and to the Santa Barbara Conference on Content and Concepts; and in 2005 to the Conference on Mental Action at the School of Advanced Studies, University of London and to my seminar at Columbia University. I have been helped by the comments of Ned Block, David Chalmers, Jerry Fodor, James Pryor, Michael Rescorla, Nathan Salmon, Stephen Schiffer, Sydney Shoemaker, Susanna Siegel and Aaron Zimmerman.

© Christopher Peacocke 2005

Gilbert Ryle once asked: ‘What is Rodin’s *Le Penseur* doing?’² My answer in this paper is that he is literally doing something, is engaged in mental actions; and our task is to say more about what this involves.

My strategy will be first to articulate some distinctive features of bodily action-awareness; then to characterize the range of mental actions; and to argue that all of these distinctive features of action-awareness in the bodily case are present also for mental actions. I will go on to consider some of the attractions and consequences of the Principal Hypothesis; to apply it in the characterization of some pathological states; and to consider some aspects of its significance for the nature of first-person thought. There are many other issues raised by the Principal Hypothesis besides those discussed here. This applies especially to its epistemological ramifications. I attempt to address them in a further paper ‘Mental Action and Self-Awareness (II)’.

I. The Distinctive Features of Action-Awareness

(a) You can be aware that you are doing something without perceiving that you are doing it. If you have had a strong injection in preparation for a root canal operation at the dentist, you may have no sensation in and around your mouth and your jaw. If you are asked to open your mouth, you can do so, and you will be aware that you are opening your mouth. This awareness exists even though you do not perceive your mouth or your lower face at all. You can be aware that you are opening your mouth without seeing or feeling your mouth, and without any of the sensations or perceptions of your own body from the inside (that is, without any proprioception). A person whose afferent nerves have been severed or have suffered decay may still be aware that he is extending his arm and pointing to the right, even though he is looking the other way, and does not perceive his own arm at all.

The same kind of action-awareness that is present in these exceptional circumstances is also something we enjoy in normal bodily action in more ordinary circumstances. Your

² ‘The Thinking of Thoughts: What is *Le Penseur* Doing?’, in his *Collected Papers vol. II, Collected Essays 1929-1968* (London: Hutchinson, 1971).

everyday awareness that you are moving your hands, turning your head, or opening your mouth is not purely perceptual.³ Even if it is true that action-awareness requires some general capacity to perceive, action-awareness on a particular occasion that you are doing something does not require you to perceive, on that occasion, that you are doing it.

(b) The content of your action-awareness is that you are doing something. It is not merely a consciousness that something is happening (though of course that is implied by the content).

This fact arguably parallels a corresponding truth about the content of perceptual states. Perception is as of states of affairs in the objective world, states of affairs of a sort that cause perceptions. If action-awareness is caused by tryings, this awareness is as of what's the case when those tryings successfully cause events in the objective world. What is then the case is that one is doing something.

(c) The content of the action-awareness is representational in the sense that in enjoying action-awareness, it seems to the subject that the world is a certain way. This seeming is belief-independent. It may seem to the unfortunate person whose arm is, unbeknownst to him, severed in a car accident that he is moving his arm, even though he has no sensation in it. This seeming has a false content. The seeming, just like a visual illusion, can persist after the subject knows his unhappy situation. In my view, action-awareness should not be identified with any kind of belief, whether first- or second-order.

Bodily action-awareness is to be distinguished from mere awareness of trying to do something. Suppose you are trying, but failing, to unscrew a tight lid on a jar. You are aware that you are trying to unscrew it. You have no awareness, either real or apparent, of the bodily action of unscrewing it. It may be that in certain circumstances, when there is no information to the contrary, tryings cause apparent action-awareness. That does not make apparent action-awareness identical with awareness of trying. It means only that what the latter kind of awareness is awareness of can itself cause apparent awareness of

³ This kind of awareness is the subject of the papers by Tony Marcel and by me in *Agency and Self-Awareness* (Oxford: OUP, 2003), ed. J. Roessler and N. Eilan.

bodily action. Apparent awareness of successfully doing something is distinct from apparent awareness of trying to do it.

Those who hold that there is nonconceptual content at the personal, conscious level will be attracted to the idea that some awareness of bodily action may have an at least partially nonconceptual content. I myself see nothing intrinsically problematic in the idea that an animal without concepts, but with nonconceptual mental representations of the world, may have a form of nonconceptual awareness of its bodily actions. The content of the awareness should be captured in a form that specifies the change in location or properties of the bodily parts that are involved in the apparent action. Such contents could be integrated into the scenario content possessed by perceptual states that I used in earlier work.⁴ Such a conception is not essential to the main theses of the present paper, however. Thinkers like John McDowell, who hold that all personal-level conscious content is conceptual, could also recognize the existence of belief-independent action-awareness.⁵ They would simply insist that its content is conceptual too.

(d) The content of the action-awareness is both first-personal and present-tensed. The content is of the form ‘I am doing such and such now’. When you take such an awareness at face value, and judge ‘I am doing such-and-such now’, your judgement is identification-free in a familiar sense. It is not the case that you are making this judgement only because, for some mode of presentation *m* other than the first person, you judge that *m* is doing such-and-such now, and you also accept that you are *m*. There are further distinctions to be drawn here, and I will return to them.

(e) (i) Action-awareness makes available demonstrative ways of thinking of particular actions. You can think of a movement demonstratively, as “this movement”, a way of thinking of a movement made available by your action-awareness of the movement.

(e)(ii) The reference of such demonstratives is determined by which movement is caused by one’s trying. It is not determined by its relations to one’s perception of the movement, if indeed any such perception exists. Nor is the reference determined by which movement

⁴ C. Peacocke, *A Study of Concepts* (Cambridge, Mass.: MIT Press, 1992)

⁵ J. McDowell, *Mind and World* (Cambridge, Mass.: Harvard University Press, 1994).

one believes it to be. There may be no such movement, even though one believes there is; or one may be wrong about which movement it is one has made. You can think to yourself, whilst making a certain gesture with your hand, ‘*This* is the victory gesture Churchill made’, where the demonstrative is made available by apparent action-awareness on this particular occasion. The demonstrative refers to the movement (type) you actually make. If your efferent nerves have been rerouted, your thought that this is the victory gesture Churchill made may be false, even though you know perfectly well which type of movement it was that Churchill used as a victory gesture. You are just wrong in thinking that *this* movement (action-demonstrative) is an instance of that movement-type you know so well.

(e)(iii) There is a distinction in the case of action-based modes of presentation which parallels that between the demonstrative and recognitional in the case of perceptual modes of presentation. The action-awareness based demonstrative ‘this movement’ requires that one enjoy at least an apparent action-awareness at the time of thinking. Otherwise it is not even available for use, just as a perceptual demonstrative is not available for use in the absence of perceptual experience. But there is also a way of thinking of a certain type of movement, made available by the fact that one can reliably make the movement. One can use this type of way of thinking even when one is not trying to make the movement so thought of. ‘I could make such-and-such gesture’, one may think, in the process of deciding how to act.

II. The nature and range of mental actions.

Events that are mental actions include instances of the following kinds:

decidings

judgings

acceptings

attendings to something or other

calculatings

reasonings

tryings.

Some types of mental event are such that instances of the type may or may not be mental actions. Such is the case with imagining. Imagining in your mind's ear Beethoven's Hammerklavier Sonata may on a particular occasion be a mental action. On another occasion, that Sonata may equally come to your auditory imagination unbidden - your imagining may be a hindrance to what you are trying to do. In this respect, imagining as a type is like the bodily type of making marks on the carpet. When someone is making marks on the carpet, that may or may not be something she is trying to do.

Within the class of mental events, what makes an event a mental action? For a mental event to be a mental action, it must consist of an event which either is, or constitutively involves, a trying. If 'constitutively involves' is allowed to count as a reflexive relation, this criterion can be simplified. To be a mental action, a mental event must constitutively involve a trying.

Every mental action involves success in something at which one may in principle fail. You may find that you cannot bring yourself to believe that p (for instance, that your friend is lying to you); you may find that you cannot bring yourself to try to do something; you may find that you cannot bring yourself even to decide to do something. Sometimes lack of success is obvious to the would-be agent himself. In other cases, an agent may have an illusion of success. A subject may think he has formed the belief that p when in fact he has not. No amount of affirming to himself that p will guarantee that he has succeeded in storing the content that p amongst his beliefs. This fact is the ground of possibility of one sort of self-deception.

The success or failure of our attempts at mental action depends upon all sorts of subpersonal conditions to which we do not have independent access, in the way in which perception gives us independent information on whether our attempts at bodily action have been successful. The real possibilities of continuing, ordinary error about some of our mental states are in this respect far more extensive than the real possibilities of such error about our bodily actions. To describe a situation in which someone is self-deceived on the issue of whether he has really unscrewed the lid off the jar would take Monty Python-like resources.

The condition I have offered for a mental event to be a mental action is the same condition as I would offer for a bodily event to be a bodily action: it must constitutively involve a trying. Mental actions and bodily actions are actions in exactly the same sense. The differences between them are the differences between the bodily and the mental.⁶

Tryings themselves featured on the above list of mental actions. This does not involve a vicious regress (nor a non-vicious one either). An unacceptable regress would be generated by the conjunction of the following propositions: tryings are actions; and for an event to be an action, it must be caused by a prior trying. That last proposition is false, however, which is why there is no regress of that sort. Tryings themselves are one of the best counterexamples to the thesis that for an event to be an action, it must be caused by a prior trying.

Though the main concern of this paper is action-awareness, the recognition that there is a range of mental action-types that includes both judgement and decision already has consequences for a range of philosophical and psychological issues, independently of theses about action-awareness. I give four examples, which should help to locate this position about mental actions in a wider philosophical and psychological landscape.

Outright judgement, something that seems not to be a matter of degree, has often seemed to play a special role in the formation of propositional attitudes.⁷ This appearance is both understandable and correct if judgements are mental actions. A mental action involves a trying, and whether you are trying to do something is not itself a matter of degree. You are either trying or you are not. What it is you are trying to do may vary in degree: you may be trying to write a long letter or a short letter, to make a lot of money or a modest sum. But whether you are trying or not is not a matter of degree. Since trying involves the occurrence of an event, an initiating event which produces an effect, it is not

⁶ I have unified mental and bodily actions by their common relation to tryings, but someone skeptical that this is the right account of action could still accept the other main claims of this paper. That skeptic could still agree that mental action awareness is a species of the same genus of action awareness that includes bodily awareness of bodily actions. The skeptic would just be offering a different account of what makes something an action, whether bodily or mental.

⁷ For one good statement of this position, see G. Harman, *Change in View: Principles of Reasoning* (Cambridge, Mass.: MIT Press, 1986), Ch. 3, in the section entitled 'All-or-Nothing Belief', pp.22-4.

surprising that it should not be a matter of degree. It is not a matter of degree whether such an initiating event occurs.

There is such a thing as trying harder or less hard to do something, and this distinction does get a grip in the mental realm as well as in the bodily. But no one who advocates the importance of degrees of belief would be tempted to identify greater degree of belief with (say) lower degree of effort in trying to make an outright judgement. Such theorists would want to contrast degrees of belief with outright judgements, however the members of each of these categories may be reached.

These points about judgement apply also to decision and to the other mental action-types. Deciding to do something cannot be a matter of degree. Again, there can be variation in degree in respect of what it is that one is deciding to do; but that is a different matter.

Not every case in which you come to believe something involves mental action. By default, we take many experiences, memories and utterances of other people at face value. What they represent as correct goes straight into our store of beliefs without any mental action. So I am not saying that every time we form a belief, even a conscious belief, there is mental action. It is however characteristic of beliefs, as opposed to more primitive representational states, that they can be assessed and reviewed. Such assessment and review does involve mental action.

A second consequence of acknowledging the existence of mental actions that include judgements concerns the idea of concepts as individuated by norms for making judgements in which those concepts are applied. These concept-individuating norms can then be seen as norms of rational action. They are norms of action applying in the special case in which the action is a mental action, that of judgement. The case is also arguably special in that the applicability of the norms is what makes something a judgement with a given content.

A third consequence concerns the philosophy of mind and action more generally, and it bears upon the existence of the phenomenon of akratic judgement. Knowing or having evidence about what it is rational to think, all things considered, or having information about what is most likely to be the case, never entails that the thinker will perform a certain *action*.

We know this very well from the case of bodily action. If judgements and decisions are mental actions, exactly the same point applies to them too. Akratic belief, and other akratic mental actions, are just as possible as akratic bodily actions. They are possible for the same reasons as in the bodily case. Mental action has all the frailties of subjection to desire, self-deception, and wishful thinking that bodily action also suffers. Mental agency is not in a privileged position vis-à-vis bodily agency. This may be humbling, but it also puts us in a much better position to explain the range of phenomena that actually occur.

A fourth consequence concerns the unified theoretical treatment of areas that have not always been considered instances of a single kind. Daniel Kahneman writes of his own and Amos Tversky's work on the two topics of intuitive thinking and of choice that it "highlights commonalities between lines of research that are usually studied separately".⁸ If both judgements and choices are mental actions, we should be ready for the possibility that, as mental actions, some of the characteristics of the mechanisms producing them are the same. In deliberating what to think, our deliberation is about a mental action, what to judge; in deliberating between options, we are deliberating about what to choose, equally a mental action.

Kahneman summarizes his views by saying that "In particular, the psychology of judgement and the psychology of choice share their basic principles and differ mainly in content." Kahneman draws a distinction between what he calls 'System 1' and 'System 2'. This distinction maps on to, and can help explain empirically, some of the distinctions I have drawn. His System 1 is "fast, parallel, automatic, effortless" and it delivers what Kahneman calls "impressions". These "impressions" are not mental actions. In this respect his comparison of them with perceptions is wholly apt. Like perceptions, they just occur to the thinker. Judgements are the output of Kahneman's System 2 and of them he writes: "In contrast, judgements are always intentional and explicit even when they are not overtly expressed" (698 ff.). This is a clear classification of judgements as actions. We will later make use of Kahneman's distinction between Systems 1 and 2.

III. The Principal Hypothesis and its Consequences.

⁸ "A Perspective on Judgement and Choice: Mapping Bounded Rationality", *American Psychologist* 58 (2003) 697-720, p.717.

The Principal Hypothesis, as I formulated it, states that a thinker's awareness of those of his mental events that are mental actions is a species of action-awareness. If mental actions are literally actions, it should not be surprising that a subject's awareness of them is of the same kind as other examples of action-awareness.

All the distinctive features of action-awareness we noted for bodily actions are also present for mental actions. We can run briefly through them, with the same lettering as above.

(a) Since you do not have perceptual experiences of your mental actions at all, and you have a distinctive awareness of them, you can certainly have this awareness without perception of them.

(b) Your awareness of your mental actions, such as your awareness that you are deciding, that you are calculating, and the like, is not merely an awareness that something is happening. It is an awareness that you are doing something, an awareness of agency from the inside.

(c) The awareness is representational: it seems to you that you are deciding, calculating, and so forth. Correspondingly there is such a thing as taking the world to be as this awareness represents it as being.

(d) The content of your awareness of your mental actions is first-personal and present-tensed: you are aware for instance that you are calculating now. An expression of this awareness with the first-person pronoun would be counted by Wittgenstein as a use of 'I' as subject. Your belief that you are calculating now does not rest on two beliefs, for some mode of presentation *m* other than the first person, that *m* is calculating now and that you are identical with *m*.

(e) (i) Mental action-awareness makes available to the thinker particular demonstrative ways of thinking of those mental actions. One can think 'this judgement', 'this calculation', and these demonstratives in thought refer to the particular mental actions awareness of which makes the demonstratives available in thought. This action-awareness makes available to a thinker ways of thinking of her own mental actions. These ways of thinking are essential to self-scrutiny and critical reflection on her own mental actions.

(e)(ii) One may have an apparent awareness of a mental action which misrepresents the mental action. When, for instance, there is a sufficiently complex structure of desires and/or emotions leading to self-deception, one may think one is judging something when one is not, and may be judging something entirely different, or nothing at all.

(e)(iii) We noted an analogue, in the case of bodily action, of the distinction between demonstrative and recognitional modes of presentation in the perceptual case. There is an corresponding distinction between two ways of thinking of mental actions. ‘This deciding’, ‘this calculation’, ‘this judgement’ are all demonstratives in thought that refer to particular mental actions. But there is also a way of thinking of a type of mental action, for example the type of being a judging that London is burning, that is individuated by its connections with one’s ability to engage in mental actions of that type. It is that way of thinking of a type that one employs when one thinks ‘If it is reported on the news that London is burning, of course I will judge that London is burning; but not otherwise’. In normal cases, when one tries to perform a mental action of this type, one succeeds. One does not normally perform mental actions of these types by doing something else. These are the analogues in the mental case of a species of basic action.⁹

IV. The Principal Hypothesis: Attractions and Possibilities

I now turn to some attractions of the Principal Hypothesis, and some theoretical possibilities and reflections that it suggests.

One of the attractions of the Principal Hypothesis is that it assimilates those conscious events that are mental actions to a wider class whose members equally share some of the distinctive features of conscious mental actions. One of the most distinctive features is that such mental actions as judging, deciding and the rest have the phenomenology of doing something, rather than involving the phenomenology of something being presented as being the case, as in perception, or as of something occurring to one, as in unintended imagination, in which cases the subject is passive. This

⁹ On basic actions, see originally A. Danto, ‘What We Can Do’, *Journal of Philosophy* 60 (63) 435-45, and for refinements, A. Goldman *A Theory of Human Action* (Princeton, NJ: Princeton University Press, 1970), Chs. 1 and 2.

active phenomenology is present for bodily action too. The action-awareness of raising one's arm is equally not that of being presented with some fact, but is rather a phenomenology of one's doing something. The position I am developing is, then, in head-on disagreement with the view that the character of conscious thought involves only states that are sensory or, like imagination, individuated by their relations to sensory states. That opposing view is well-formulated (though not fully endorsed) by Jesse Prinz, who writes "When we introspect during thought, all we find are mental images, including auditory images of natural-language sentences (subvocal speech). With no phenomenal traces of nonsensory representations, it is tempting to conclude that all thought is couched in perceptual imagery".¹⁰

Here are three cases, subjectively different, in which exactly the same words – for instance, 'Meeting tomorrow!' - may occur in your mind's ear:

- (a) The words may just passively occur to you; this could be memory or unbidden imagination.
- (b) You may be judging that the meeting is tomorrow, on the basis of remembered evidence.
- (c) You may be making a decision to convene the meeting tomorrow.

The difference between these three – imagining or remembering, judging and deciding - is certainly not something within the phenomenology of passive imagination or presentation. Nonetheless, it is a feature of your consciousness that you are, for instance, judging something rather than forming an intention. It is equally a feature of your consciousness if you are merely passive in this respect. Action-awareness is given as action-awareness, and is subjectively different from merely passive states. Any description of your conscious state is incomplete if it omits the characteristics of action-awareness. "No difference in imagistic or presentational phenomenology" does not imply "No difference in phenomenology at all".

¹⁰ *Furnishing the Mind: Concepts and their Perceptual Basis* (Cambridge, Mass.; MIT Press, 2002), p.103.

The point applies even when there are no mental images or perceptions involved at all. Someone, Rodin's *Penseur* with his eyes closed, may be passively drifting in thought, and nothing may come to his mind; or he may be thinking hard about how to solve some theoretical or practical problem - and equally nothing may come to mind. These are very different total subjective states. The person who is concentrating on finding a solution to a problem is actively trying to do something in thought; and this contributes to the phenomenology of his state.

Correspondingly there is a difference in imagining being in these two states. This is what one would expect if imagining, from the inside, being in a certain state is subjectively imagining what it is like to be in that state. Imagining drifting aimlessly in thought is different from imagining concentrating on solving a problem.

Recognition that there is a distinctive category of mental action-awareness can account for many of the features of conscious thought that so engaged Gilbert Ryle in his late writings on the topic.¹¹ Someone so inclined could devote a whole paper (or more) to this topic. Here I just give two examples.

Ryle repeatedly emphasized that neither the occurrence of any one particular event involving the imagined uttering of words, or visualizing of scenes, or anything else of the sort, or any disjunction thereof, is what constitutes judging, when out on a drive, that the petrol (gas) station at the next village may be closed for Sunday.¹² In my view, Ryle is right about this. Under the Principal Hypothesis, his point is just what one would expect. None of the things Ryle rightly cites as insufficient for judgement involves action-awareness of judging, which is something additional to, and not ensured by, any amount of word-imagining, picturing in one's mind's eye, and the like.

The other example involves Ryle's long-standing (perhaps even fatal) attraction to 'adverbial' theories of mental phenomena. He notes that in the case of bodily events, some of them have 'thick' as well as thin descriptions. His example is that a hitting of a ball with a golf club may also be a 'practice approach shot', and 'a piece of self-training' (474). He says these thick descriptions involve 'intention-parasitism', and that the same

¹¹ See especially the papers 'A Puzzling Element in the Notion of Thinking', 'Thinking and Reflecting' and 'The Thinking of Thoughts' in volume II of his *Collected Papers*.

¹² See 'A Puzzling Element', p.393 ff..

phenomenon is found amongst mental events, which may, in the case of a composer, be tryings-out, modifications, assemblies, and in the case of other projects in thought, may be serving many other purposes. He rightly concludes that descriptions of mental events involving intentionality on the part of the thinker will not be determined by neutral characterizations of the subjective contents of imaginings and visualizings; and that the intentional characterizations may be correct for many different kinds of imaginings and visualizings (476-9). What Ryle call intention-parasitism is possible only where there is mental action. From the standpoint of the present paper, there is nothing either adverbial (or higher-order, for that matter) in a mental event's being a mental action. To be a mental action, the event must have the additional property of having been produced in the right way by the subject of the event. When it is so, it is then possible for the 'thick' descriptions that Ryle mentions to get a grip. (The 'intention-parasitism', insofar as I understand it, is also not necessary for an event to be a mental action: I may actively imagine the Hammerklavier on a whim, and not in pursuit of some further purpose.)

In current philosophy of mind, there is a range of kind of states each of which is recognized as having representational content, in the sense that in being in one of these states, it thereby seems to the subject as if that content is correct. This seeming may be overruled by judgement, or it may be taken at face value. In either case, the state's possession of a representational content should not be identified with the subject's judging that content (or a corresponding content) to be correct. States currently recognized to possess such representational content include at least the following three kinds. There are perceptual states, in which, in having an experience in a particular sense modality, it seems to the subject that the world is a certain way. There are states of pure thought, in which it strikes one as the case that (say) the American Declaration of Independence was signed in 1776, where this purely propositional impression does not need to correspond to any personal memory. There are representational states of personal memory, in which one has a memory of, say, walking on the beach at Big Sur, and it thereby seems to one that one was there. To this list of kinds I suggest that we should add action-awareness. Your apparent action-awarenesses of raising your arm, of judging that it is time to leave, of calculating the sum of two numbers each represent you as doing these very things. And just as a memory-impression may be a memory of your perceiving

something in the past, and that represents you as so perceiving, a memory may also be of your doing something, and represent you as having done that thing. A memory of walking along a beach will commonly do both. To give a correct account of the relation between these states and the kinds of content they can contain, and to do so in way that provides a philosophical resource, is a general challenge. I will return to it in the particular case of action-awareness and its contents.

When a subject has an action-awareness that he is φ -ing, for example, that he is turning the left-hand knob, all the contents of the that-clause contribute to the character of his awareness. There is the action-type of turning, different from that of pushing or pulling, and which he is aware of performing. Similarly action-awareness of judging is different from action-awareness of coming to a decision. But the intentional objects of the action also contribute to the awareness too. One is aware that one is turning this knob rather than that one (both demonstratively given in thought). Similarly one is aware that one is judging one complete propositional intentional content rather than another; and that one is coming to one decision rather than another.

In earlier writing I drew a distinction between being the object of attention, and occupying attention.¹³ In conscious thought, your attention is occupied, but there need not be anything which is the object or event to which you are attending (not even an apparent object). The Principal Hypothesis contributes to an explanation of this difference. In ordinary action-awareness of bodily action, such as your awareness of raising your arm, your action-awareness need not involve your attending to your arm, or to its rising, even though your conscious action can certainly occupy your attention. If conscious thought is action-awareness, we would expect the same. The action of which you are aware in a distinctive way – making a judgement, forming an intention - does not involve the making of the judgement, or the formation of the intention, being the object of your attention. Rather, as in the case of bodily action, making the judgement, or forming the intention, *occupies* your attention.

V. Describing and Explaining Schizophrenic Experience

¹³ See my contribution to *Knowing Our Own Minds* (Oxford: OUP, 1998), ed. C. Wright, B. Smith, C. MacDonald.

Our Principal Hypothesis states that a thinker's awareness of those of his mental events that are mental actions is a species of action-awareness, with all the distinctive characteristics of action-awareness. The Principal Hypothesis bears upon our understanding of the phenomenon of schizophrenia. The Hypothesis contributes to a correct characterization of what it is that the schizophrenic subject lacks. It is equally essential to providing a deeper unification of some of the symptoms of schizophrenia. The distinctions drawn upon in elaborating the Hypothesis are also important to current psychological theories in their explanation of the occurrence of schizophrenia.

I divide the significance of the Principal Hypothesis for schizophrenia into five different headings.

(i) What the schizophrenic subject lacks in the area of conscious thought is action-awareness of the thoughts that occur to him. To enjoy action-awareness of a particular event of thinking is to be aware, non-perceptually, of that thinking as something one is doing oneself. The awareness of one's own agency that exists in normal subjects is missing in, for example, the schizophrenic experience of 'thought insertion'. One schizophrenic subject famously reported: "The thoughts of Eamonn Andrews [a UK television presenter in the 1960s] come into my mine. He treats my mind like a screen and flashes his thoughts on to it like you flash a picture"¹⁴.

It is important to characterize the schizophrenic's consciousness as lacking action-awareness. It is not merely that these subject report that their conscious mental events are caused by external, intervening agents. Even when they no longer report that they are so caused, because they are persuaded of the non-veridicality of these conscious events, these subjects' experience of passivity persists. Action-awareness is still absent, whatever the schizophrenic subject's own beliefs, if any, about why he is having he is having mental events from which the action-awareness is absent. Precisely because action-awareness is, like perception, belief-independent, it cannot be restored simply by altering someone's beliefs.

¹⁴ *Schizophrenia: A Very Short Introduction*, by C. Frith and F. Johnstone (Oxford: Oxford University Press, 2003), p.36.

The schizophrenic condition is also sometimes characterized as a ‘failure to distinguish between ideas and impulses arising from within the subject’s own mind and perceptions arising from stimuli in the external world’ (ibid. 37). But subjects do draw the distinction. The ability to draw the distinction is implied by the subject’s own description of thought-insertion just quoted. It is in part because the distinction is drawn that the conscious states of schizophrenia are so alarming to their unfortunate subjects. The right way to formulate the point about the distinction rather involves action-awareness. The schizophrenic subjects lack the action-awareness in thought present in normal subjects, an awareness that, in its representational content, draws the distinction between events produced by oneself and events produced by others in the right place.

(ii) The Principal Hypothesis provides a straightforward unification of some of the symptoms of schizophrenia in thought and some of its symptoms in bodily action. Some schizophrenic subjects experience delusions of control of their body by an external agency. “It is my hand and arm that move, and my fingers pick up the pen, but I don’t control them. What they do is nothing to do with me.” (ibid., 36). Sean Spence asked subjects with delusions of control to perform a simple bodily task of holding a lever and producing a random sequence of movements. They performed this task normally, but still reported that their movements were controlled by alien forces (ibid. 137).¹⁵

The Principal Hypothesis states that awareness of mental actions is action-awareness of the same sort as occurs in bodily-action awareness. Subjects who lack action-awareness of the thoughts they are in fact producing must have some kind of impairment of the mechanism that, in healthy subjects, produces action-awareness. But if action-awareness in the bodily case is awareness of exactly the same kind as in the case of conscious mental actions, it is to be expected that some cases of impairment of the mechanism producing action-awareness would affect awareness of bodily actions too. This is just what one finds. Symptoms that might otherwise seem somewhat diverse, and might even raise doubts about whether there is a single underlying condition of which they are both manifestations, are in fact unified by the Principal Hypothesis.

¹⁵ SA Spence, DJ Brooks, SR Hirsch, PF Liddle, J Meehan, and PM Grasby, ‘A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control)’ *Brain* 120 (1997) 1997-2011.

(iii) Some of the phenomena of schizophrenia highlight, and cannot be properly characterized without, the distinction between action-awareness and awareness of goals and intentions. The idea of a defect in awareness of goals and intentions has sometimes played a large role in some earlier theorists' explanation of schizophrenia. It played such a role in Christopher Frith's 1992 account in *The Cognitive Neuropsychology of Schizophrenia*.¹⁶ But we should remember the subjects in the Spence study just mentioned, in which subjects succeeded at simply bodily tasks they were instructed to carry out, but who still experienced delusions of control. These subjects knew their goal and their intention – it was to perform the task the experimenter had requested. Their abnormality is not in failing to represent their goal or intention correctly, but in their lack of action-awareness of their bodily actions as their own.

(iv) There is a theory proposed by Irwin Feinberg, and developed further by Frith and Johnstone, which proposes for schizophrenia an analogue of Helmholtz's famous 'corollary discharge' in visual perception.¹⁷ Helmholtz offered an explanation of why the world does not seem to move when you move your eyes, even though the image of objects moves on the retina as your eye moves. According to Helmholtz, just prior to a movement of your eyes there is a corollary discharge caused by the attempt to move the eyes, and this discharge permits a computation of the location of objects in the environment that takes into account the movement of the eyes. Frith and Johnstone write that "Patients with delusions of control and related symptoms have problems that suggest that they cannot monitor their own movements in the normal way" (133). When we regard consciousness of mental actions as a species of action-awareness, such awareness can be accounted for in this explanatory structure. The natural conjecture, given all the evidence to date, is that:

(a) When there is no corollary discharge, there is no action-awareness of the movement in question as one of your own actions, and this applies quite generally, both in bodily and mental cases. If the corollary discharge theory is correct, this hypothesis would explain

¹⁶ Erlbaum: Hove, East Sussex, 1992. See the summary pp.133-4, and earlier in the same chapter.

¹⁷ See Frith and Johnstone; I. Feinberg, 'Efference copy and corollary discharge: implication for thinking and its disorders', *Schizophrenia Bulletin* 4 (1978) 636-40; H. von Helmholtz, *Treatise on Physiological Optics* (New York: Dover, 1962).

the absence of action-awareness in schizophrenic subjects, again both in bodily and in mental cases.

(b) If the corollary discharge is caused by trying to perform the action in question, in normal subjects, that explains why, when there is no evidence to the contrary, trying itself causes an (apparent) action-awareness. Computationally, it is for the agent exactly as one would expect it to be when there is action. This would also explain the apparent action-awareness in trying to move a severed limb. It may also explain some illusions of having formed a belief.

(v) There is a syndrome of symptoms in schizophrenia having to do with a loss of will, an absence of spontaneous action and thought, and blunted emotional responses. Action-awareness is the most obvious and fundamental manifestation in conscious life of oneself as a successful agent. When this awareness is lacking, it is not surprising that a subject's sense of himself as an agent should suffer, and that he should be less motivated to action and spontaneity. When your actions, however extensive, are experienced only passively, it is hard to conceive of yourself as a successful agent. Absence of action-awareness is not an isolated phenomenon of consciousness, but has ramifying effects, both for the emotions of the schizophrenic subject and for his self-conception.

Obviously there is much about schizophrenia that the Principal Hypothesis does not explain. A full understanding has to explain the prevalence of the impression of control by alien agencies and forces. Why an absence of action-awareness should lead to this specific kind of illusion needs an empirical explanation by resources going far beyond those of the Principal Hypothesis. My position is only that we need the distinctions I have been drawing to characterize and unify the schizophrenic phenomena. We will not have a proper empirical explanation of the phenomena without an accurate characterization of what it is that has to be explained.

VI. The First Person in Action Self-Ascriptions

I now turn to the role of the first person in action-awareness. I define a use, on a particular occasion, of the first person in thought as a *use of 'I' as agent* as one in which that use occurs in a first-person judgement made simply by taking the representational

content of an apparent action-awareness at face value. The uses of the first person in the judgements ‘I am pressing the button’ and ‘I judge that Bush will be re-elected’ will be uses of ‘I’ as agent when made by taking the corresponding action-awarenesses at face value.

Uses of ‘I’ as agent are uses of ‘I’ as subject, in the sense employed by Wittgenstein in *The Blue and Brown Books*, and later so well-elucidated in Sydney Shoemaker’s important papers.¹⁸ As we noted, in ordinary circumstances, when a thinker uses ‘I’ as agent in a judgement ‘I am ϕ -ing’, his judgement does not rest on a pair of beliefs that m is ϕ -ing, for some m distinct from the first person, together with an identity belief ‘I am m ’. I do not have first to judge ‘that person is pressing the button’, or ‘CP is pressing the button’, before I am in a position to judge ‘I am pressing the button’. Action-awareness already has a first-person component in its intentional content. If the thinker is taking that awareness at face value, no such identity belief is needed for the thinker to be in a position to make a self-ascription of the action in question. The case quite unlike that in which my belief ‘My car alarm is sounding’ is based on the two beliefs ‘That car’s alarm is sounding’ and ‘I am the owner of that car’. In Shoemaker’s terminology, judgements ‘I am ϕ -ing’ involving the use of ‘I’ as agent are immune to error through misidentification relative to (the first occurrence of) the first person.

For enthusiasts about these distinctions, this is arguably a case of what Shoemaker calls *de facto* immunity (in ‘Persons and their Pasts’). In a world in which devices or Wilder Penfield-like persons intervene after one’s tryings, and, by means of some randomizing mechanism, may or may not make their intended bodily and mental effects come about, there could regularly be incorrect apparent action-awarenesses. In such a world thinker could introduce a demonstrative ‘That_A agent’, that refers to whoever is the agent of the event of which the subject has a token action-awareness A. This is the action-analogue of the demonstratives for times and places I imagined in *Sense and Content* for in cases in which there massive time-lags in perception, or perceptions as from places other than one’s current location. [ref] In those circumstances, one could

¹⁸ ‘Self-Reference and Self-Awareness’, and ‘Persons and their Pasts’, both reprinted in his *Identity, Cause and Mind: Philosophical Essays* (Cambridge: Cambridge University Press, 1984).

reasonably wonder ‘Am I identical with that _A agent?’. But this is no more our actual situation with respect to agency than is the corresponding situation for the invented temporal and spatial demonstratives. Whatever the correct explanation of the phenomenon, contingent features of our actual circumstances can have a bearing on what is required for coming to make a judgement reasonably.

The existence of a use of ‘I’ as agent and the nature of the conscious states on which these uses are based can help explain some of the illusions, in the history of philosophy, to the effect that there exists a transcendent subject whose transcendent operations affect the spatial world, and the mental world. In the apparent action-awareness ‘I am ϕ -ing’ itself, the subject is not given as having a location in the spatial world, nor as having spatial or material properties. This applies to predications of bodily actions of ϕ -ing, as well as to mental actions. The apparent action is bodily, but the subject who is represented as doing it is not represented in the awareness as a spatial object, or as having spatial properties, itself.

It would be a terrible fallacy – one of those nonsequiturs of ‘numbing grossness’ – to conclude from this fact that the subject referred to in such thoughts and awarenesses does not have a spatio-temporal location and does not have spatio-temporal and material properties. It would be a fallacy even to conclude that the subject referred to does not need to have such properties. But it would be a brave person who, on reading the works of those who have postulated a transcendent subject, concludes that no such fallacious transition is hovering over their writings. This is particularly so in the case of those writers who have placed some species of agency in a noumenal realm.

As is often the case with the postulation of transcendent subject-matters, the motivation for the postulation involves a genuine insight, misapplied. It is right to hold that much thought is mental action, and so must be explained in the same general way that other action is. It is wrong to think that a transcendent subject is either necessary or possible in explaining these distinctive phenomena.

The case of action-awareness is a distinctive one amongst the range of phenomena that can generate illusions of transcendence, in that the intentional content of the awareness itself contains the first person. A wide range of other cases that generate the

illusion have the property that in *Being Known* I called ‘representational independence’.¹⁹ When self-ascribing a perception, or an occurrence of a passive occurrence of a conscious thought to oneself, one does not rely on a conscious state which represents oneself as enjoying that mental state. Rather, one moves rationally from that mental state itself to a self-ascription. There are thus two rather different ways in which it may come to seem that ‘I’ refers to something without spatial or material properties. One way is for the transition to a judgement to move from a state which does not contain the first person in its intentional content (or not as standing in the relation self-ascribed). The other is for the rationalizing state to contain the first person in its intentional content, but for that content not to represent the subject as having spatial and material properties. Described in the abstract, this case might seem to be of questionable possibility; but it is this possibility that action-awareness realizes. (It also follows that a different explanation of the entitlement to the transition must be given in the case of action-awareness than in the representationally-independent cases. In *Being Known*, I offered what I called ‘the delta account’ (Sections 6.2, 6.3). The account above of entitlement for the case of action-awareness, which in abstract structure more closely parallels that for perceptual judgements, is quite different from the delta account.)

While all uses of ‘I’ as agent are uses of ‘I’ as subject in our ordinary circumstances, the converse is not true. There are uses of ‘I’ as subject, even uses in the self-ascription of attitudes, that are not uses of ‘I’ as agent. An example of Richard Moran’s illustrates the possibility.²⁰ You may come to the conclusion that you believe that someone has betrayed you on the basis of information about your feelings, emotions, and other judgements. As Moran writes, “insofar as it is possible for one to adopt an empirical or explanatory stance on one’s own beliefs, and thus to bracket the issue of what their possession commits one to, it will be possible for one to adopt this stance to anything theoretically knowable, including private events or attitudes that one may be somehow aware of immediately, without inference” (92). “We may allow any manner of inner events of consciousness, any exclusivity and privacy, any degree of privilege and

¹⁹ C. Peacocke, *Being Known* (Oxford: Oxford University Press, 1999), section 6.1.

²⁰ See R. Moran, *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton: Princeton University Press, 2001).

special reliability, and their combination would not add up to the ordinary capacity for self-knowledge.” (93). Suppose you come to the conclusion that you believe that a certain person has betrayed you, and your evidence for this self-ascription consists of your other mental states that, in self-ascribing, you use ‘I’ as subject. The evidence might, for instance, include your emotions of anger or irritation at the person, and your self-ascriptions can involve uses of ‘I’ as subject. Your inferential judgement ‘I believe that person has betrayed me’ would, in these circumstances, not be reached by some identity inference from two premises of the form ‘*m* believes that that person has betrayed him’ and ‘I am *m*’. So the self-ascription does, in ordinary circumstances, involve a use of ‘I’ as subject. But it is not a use of ‘I’ as agent based on an action-awareness of judging that that person has betrayed you. There is, in the example, no such action-awareness, and no such judgement for there to be an action-awareness of.

In this example, the self-ascription in ‘I believe that that person has betrayed me’ uses ‘I’ as subject because the premises from which it is reached also use ‘I’ as subject. But the same propositional evidence about some person given in a third-person way *m* could equally, and in normal circumstances, support the conclusion ‘*m* believes that that person has betrayed him’. By contrast, when one self-ascribes a belief on the basis of action-awareness, such awareness involves the first-person essentially. Reliance on action-awareness is a way of coming to ascribe an attitude that one can, in ordinary circumstances, use only in ascribing attitudes to oneself. In this respect, it is unique to the first person.

VII Concluding Remarks: Rational Agency and Action Awareness

Rational agency and action awareness are coordinate elements in being a rational subject. Neither element seems to be definable in terms of features of the other.

The idea that the nature of action awareness is explicable without reference to rational agency is immediately puzzling. As we emphasized, an apparent action-awareness has a representational content whose correctness requires that the subject of the awareness be the agent of the event which the awareness represents the subject as producing. The correctness of the apparent awareness requires rational agency. If the

apparent action awareness is correct, there will be rational agency. Further, if the apparent awareness is apparent awareness of some state of affairs whose existence is independent of the apparent awareness, as it seems to be, the prospects for reducing rational agency to features of action awareness are poor.

What of the converse direction? Can action awareness be reduced to other features of rational agency? Perhaps the most salient candidate for reduction is a philosophical explanation of facts about action awareness in terms of a thinker's knowledge of his intentions. Is a thinker's knowledge of what he is doing really explained by his knowledge of his intentions in acting?

There are at least two problems with this idea. The first is that a thinker can intend to act at a given time; may know that that time is now; but may yet fail even to try to act. When the thinker does try to act, how does he know that he is trying? It is no defense of this position to say that he is aware that he is trying. Trying itself is a mental action, and awareness of it is a case of action awareness, the phenomenon that this account was trying to explain in terms of knowledge of intentions.

The other problem is that even if we grant that the subject knows that he is acting, his knowing that he is intending to ϕ does not imply that he has an action awareness of ϕ -ing. In operating the photocopying machine, I know that I am intending to make a good copy of a document. I do not have an action awareness of making a good copy of the document. I have only an action awareness of moving my hands and, possibly, of pressing the button. Only by opening the lid of the machine and perceiving the result do I become aware that I have made a good copy, if I have. The same applies even when it is not a question of operating machines whose results are not immediately open to view. If I am novice at Greek, then whether I have successfully written a Greek letter zeta, or traced its shape correctly in the air, may not be something I know or am aware of simply by having an operative intention to do so. It is not true that you know what you are really doing simply by knowing your intentions in acting.²¹

²¹ Some of these cases appear to be counterexamples to what Richard Moran calls 'Anscombe's Condition': "If he can only know what he is doing by observing himself, that would be because, described in *these* terms (e.g. clicking out the rhythm [while pumping water]) his action is *not* determined by his primary reason, is not undertaken by him as the pursuit of some aim. Otherwise, he would know what he is doing in knowing

As we stand back from the details of these issues, the deep question that emerges is why there is a connection between rational agency and awareness. It is an instance of a more general connection of which we need a better understanding. In the case of the non-mental world, we know that a rational subject can judge and act only on what he is aware of. We do not expect the informational states of the blindsight subject, however reliable, to explain his rational decisions and actions. If they do explain his decisions and actions, it is not by rational transitions of thought. What applies to the non-mental world holds here equally for the mental world. A rational subject can make decisions and mental self-ascriptions, and keep track of his own mental events and states, only if he is aware of them. The awareness may be of a distinctive kind, as I have been arguing that it is, but the general principle still holds. Further investigation of this territory should include exploration and explanation of the internal connections between awareness and the rationality of thinkers.

his practical reasons for adopting this aim.” See his *Authority and Estrangement: An Essay on Self-Knowledge* (Princeton, NJ: Princeton University Press, 2001), pp. 126-7. My own view is that Moran’s fundamental insights on the role of agency in a range of cases of self-knowledge can be reconciled with, and may be strengthened by, an account of action awareness in self-knowledge. The relation of Moran’s fertile discussion to the theses of the present paper merits extensive independent consideration.