For the volume *Mental Action*, ed. L. O'Brien and M. Soteriou, forthcoming from Oxford University Press (2009).

Mental Action and Self-Awareness (II): Epistemology[1]

Christopher Peacocke

We often know what we are judging, what we are deciding, what problem we are trying to solve. We know not only the contents of our judgements, decidings and tryings; we also know that it is judgement, decision and attempted problem-solving in which we are engaged. How do we know these things?

Such pieces of knowledge are members of the wider category of knowledge of our own mental actions. My aim in this paper is to give a philosophical account of the nature of our knowledge of our own mental actions. Any account of this knowledge has to dovetail with a theory of the nature of mental action itself. The account of mental action on which I will be drawing for this purpose is one that endorses these principles:

> Mental action is a genuine subspecies of action in general. The differences between mental action and bodily action are fundamentally only the differences between the mental and the bodily.
>
> You can be aware that you are performing a certain action without perceiving that action, and without bodily perception from the inside of the motions involved in the action. This distinctive action awareness exists for mental action as well as for bodily actions. Perhaps someone will insist that action awareness is simply another form of perception, a form we should recognize in its own right. I would dispute that, but for present purposes we do not need to enter that discussion. What ought to be uncontroversial, and all that matters for the position I will be developing, is that you can be aware that you are doing something without

[1] I thank Lucy O'Brien, Matthew Soteriou and two anonymous referees for OUP for helpful comments.

perceiving your action in any of the ordinarily recognized senses of vision, touch, proprioception, hearing, taste or smell. You can be aware that you are raising your fully anaesthetized arm without any feeling in the arm, and whilst looking the other direction.[2]

The content of such action awareness is first-personal and present-tensed. It has the form 'I am doing such-and-such now'.

This conception of mental action is one for which I argued in an earlier paper, 'Mental Action and Self-Awareness (I)'.[3]

We need to subdivide the epistemological issues about knowledge of our own mental actions. A first set of questions concerns the nature of the way (if any) in which we come to know what mental actions we are engaged in. What is this way, and what gives it the status of a way of gaining knowledge?

Once we have proposals for answering these initial questions, we have to engage them with some issues that have been central in recent debates about self-knowledge. Many recent treatments of self-knowledge have, rightly in my view, rejected perceptual models of self-knowledge, for a variety of reasons. The question arises: do the proposals I offer fall to the same objections as those to perceptual models of self-knowledge, and if not, why not?

Another fundamental and continuing contemporary issue I will address is that of the consistency of our distinctive self-knowledge with externalism about intentional content. The account of knowledge of our own mental actions I offer makes possible an answer that addresses the concerns of those who think that certain kinds of account of

[2] The question of what it is for an event to be a perceptual experience is interesting, potentially significant, and, to the best of my knowledge, underdiscussed. One difference between action awareness and perceptual awareness is that action awareness does not involve sensational properties in the way all genuine perceptual experience does. (There is some preliminary discussion in my paper 'Sensational Properties: Theses to Accept and Theses to Reject', forthcoming in a special issue of the *Revue Internationale de Philosophie*. ed. J. Proust.) Another difference is that perceptual awareness provides objects of attention, and action awareness does not. It is a question whether these differences are fundamental, or are rather by-products of something else that is more fundamental. As these tentative remarks suggest, the whole issue merits further consideration.

[3] In *Contemporary Debates in the Philosophy of Mind* ed. J. Cohen and B. McLaughlin (Oxford: Blackwell, 2007).

this distinctive self-knowledge cannot be reconciled with the externalist characteristics of intentional content. These are by no means all the issues that arise in the epistemology of mental action; but any account that cannot address these issues satisfactorily would be a non-starter. I take them in turn.

I What, if anything, is the way in which we come to know of our own mental actions?

The distinctive way in which a subject comes to know of his own mental actions is by taking an apparent action awareness at face value. You judge that it will rain. When so judging, you have an apparent action awareness of your judging that it will rain. It seems to you that you are judging that it will rain. By taking this awareness at face value, you come to know that you judge that it will rain. In another case, you may have an apparent action awareness of your calculating the sum of two numbers; by taking this awareness at face value, you come to know that you are engaged in calculating the sum of two numbers; and so forth.

Apparent action awareness is a belief-independent event. A thinker may or may not endorse in judgement the content of an apparent action awareness. An amputee may know very well that if he tries to raise his missing right arm, he will have an apparent action awareness of raising it, even though he is not in fact raising it. This subject will not endorse the content of his apparent action awareness. Because action awareness is not the same as judgement or belief, a self-ascription of an action made by taking an apparent action awareness at face value is not reached by inference. It is no more inferential than is a perceptual judgement made by taking a perceptual experience at face value.

Because action awareness is not judgement or belief, a self-ascription of a mental action made by taking an action awareness at face value is a counterexample to the principle that knowledgeable mental self-ascriptions must be made by observation, by inference, or by nothing.[4] On the present view, that is a spurious trilemma. Action awareness is not perception, and can exist in the absence of perception of the action of

[4] The first sharp formulation of this view known to me is in P. Boghossian, 'Content and Self-Knowledge', *Philosophical Topics* 17 (1989) 5-26, at p.5.

which it is awareness. Judgements based on action awareness are not reached by inference, since action awareness is not judgement or belief. And judgements based on action awareness are not based on nothing, since action awareness is a real state of consciousness, available for rationalizing certain judgements. (These points apply equally to self-ascriptions of bodily actions too. The trilemma "by observation, by inference, or by nothing" is similarly inapplicable to knowledge of one's bodily actions, when based on action awareness.) We should draw the conclusion that the model of observation is not the only model available for a substantive, non-inferential epistemology of first-person mental ascriptions.

As in other cases in which the content of a belief-independent state is taken at face value, there are (at least) two ways of developing an account of the matter. On one approach, the content of the belief-independent state does not involve conceptualization of the action-type or the content of the mental action. This is the option that holds that the content of the apparent action awareness is nonconceptual representational content. A theorist developing this treatment will be likely to hold that part of what it is to possess the concept of (say) judgement is to be willing apply the concept rationally to oneself in response to an apparent action awareness of one's judging a certain content. So a concept of an action-type, judgement, is individuated in part by its relations to action awareness of judgement; just as, on the perception side, an observational shape concept (say) is individuated in part by its relations to perceptual content with a certain nonconceptual representational content. On a different, conceptualist, approach, in the style of McDowell, it would be maintained that all personal-level content is conceptual, and this applies as much to the content of apparent action awareness as, on the conceptualist view, it applies to perceptual content.[5] On this view there would be no such thing as conscious action awareness without conceptualization of what it is awareness of.

The issues at stake in choosing between these two approaches are well-known from the case of perceptual content, and I will not pursue them at this point. As far as I can see, the thesis that we come to know of our mental actions by our action awareness of them is neutral between these opposing lines of thought. The thesis can consistently be accepted by the believer in nonconceptual content, and can consistently be accepted by

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

his conceptualist opponent. In this area, there may seem to be a special problem for the friend of nonconceptual content. If the content of a judgement or decision, say, is conceptual, as it is, how can the action awareness of making the judgement or decision be nonconceptual? Here we must distinguish what the awareness is of at the level of reference, and how events, things and properties at the level of reference are given in consciousness. A characterization of a state's content as nonconceptual has to do with how things are given, not which things are given. A state of consciousness can have a nonconceptual content concerning things that include concepts. This is something we should already recognize independently to be possible if we grant that there can be conscious thinking by children who do not have concepts of concepts and do not have concepts of intentional contents built up from concepts. It is one thing to be employing concepts, and have conscious states whose content involves those concepts. It is a further thing to conceptualize those intentional contents themselves.

Is action awareness philosophically explicable in terms that do not involve reference to subjective, conscious states and events? I call the claim that it is so explicable 'The Reducibility Thesis'. Under the Reducibility Thesis, however it is developed, action awareness is not something fundamental, and to understand the role of action awareness in our thought we must look to more fundamental conditions that do not involve consciousness. Any epistemological role played by action awareness would then be played by these more fundamental conditions not involving consciousness. But I dispute the Reducibility Thesis.

How might the Reducibility Thesis be developed? Can we say that action awareness consists in no more than an action's being a result of the operation of rational agency? That would need qualification on several fronts. (a) Making photocopies is an action of mine, but I need not have an action awareness that I am making copies. My action awareness is of pressing certain buttons on the machine. To accommodate this, the Reducibility Thesis could be confined to types of action that are basic for the agent, actions the agent does not, in the content of his intentions, do by doing something else. The defender of the Reducibility Thesis would need to make this restriction to basic action-types both for bodily actions and for mental actions. (b) The Reducibility Thesis would also have to make some accommodation of what Brian O'Shaughnessy calls sub-

intentional acts.[6] Tapping your toes, moving your tongue are actions. You can become aware of them, and indeed come to have a distinctive action awareness of them, but it is not clear that that action awareness was already there when the actions were first performed. The defender of the Reducibility Thesis may make various moves at this point. One would be to insist that there is action awareness even in these cases, but its content does not go even into short-term memory. Another would be to hold that the Reducibility Thesis holds only for the fully intentional acts of a rational agent. Both of these responses would need some work to become convincing; but let us leave speculation on how that might be done, because there is a deeper, and quite general, problem for the Reducibility Thesis.

It seems there could exist a being whose movements and whose changes in mental state are sensitive to the content of its beliefs and intentions, but whose tryings and actions, both bodily and mental - if actions they be - do not involve any action awareness, either real or apparent. These beings would have to perceive their bodily actions, through vision, touch or proprioception to know that they are occurring. Would such subjects be exercising rational agency as that notion is understood within the terms of the Reducibility Thesis? If so, then the notion of rational agency employed in the Reducibility Thesis is so thin that it seems incapable of capturing action awareness at all. But if such subjects are not so conceived of possessing rational agency, it seems that action awareness, of both bodily and mental actions, has to be conceived as a co-ordinate element in rational agency in its own right. An explanation of the epistemology of action, both bodily and mental, has to go beyond materials that could equally be present in cases that are wholly non-conscious on the action side.

This is a point that bears not only on the Reducibility Thesis, but equally on any attempt to explain certain kinds of self-knowledge in terms of agency alone. We need to recognize a co-ordinate, and irreducible, element of consciousness in rational agency and action awareness as we actually have it. In my judgement, many illuminating recent discussions of self-knowledge and agency work only because there is a background assumption that we have action awareness of our bodily and mental actions.

---

[6] B. O'Shaughnessy, *The Will* (Cambridge, England: Cambridge University Press, 1980), Volume 2, Chapter 10, 'The sub-intentional act'.

A view that is distinct from the Reducibility Thesis, but which still moves in a deflationary direction, is that though there is such a thing as action awareness, it is really no more than is involved in the wider species of non-inferential knowledge. A person can know that Beethoven was born in Bonn, and this knowledge need not be based on anything more than the propositional impression, delivered by memory, that Beethoven was born in Bonn. Is action awareness any different from the mere propositional impressions involved in such non-inferential knowledge? In fact the mere propositional impression that one is acting in a particular way, whatever its source, is something weaker than action awareness, and is not sufficient for the distinctive phenomenology of action awareness. One symptom of this difference is that action awareness makes available a distinctive variety of demonstrative ways of thinking of actions given in one's action awareness. You have an action awareness of *this* raising of your arm. There would not even exist a distinctive demonstrative *this raising* unless you had this action awareness. If action awarenesses, real or apparent, were mere propositional impressions, it would be unintelligible how the conceptual component *this raising* could exist to be available for use in your thought, for no account would be available of how its reference is determined, in the absence of real or apparent action awareness. Though action awareness is not perceptual awareness, the problems for such a position are structurally entirely analogous to those attending a position that aims to reduce perceptual experiences to mere propositional impressions. Those propositional impressions, in the perceptual case, would have to contain perceptual demonstratives such as *that cup*, *that door*, if the account is to get off the ground at all. But just as in the action awareness case, these demonstratives are individuated by their relations to perceptual experience. Mere propositional impressions, both in the action and in the perception cases, are inadequate to a description of these phenomena.

II What makes the way identified in the action awareness account a way of coming to *know*?

It is widely accepted that there is a range of observational concepts – concepts of shape, size, orientation, colour, texture, amongst others – that are individuated in part by the fact that certain perceptual experiences give reasons to apply these concepts to objects or events presented in those perceptual experiences.[7] Theorists differ on how this individuation works, but there is less disagreement that there is some such individuative link between these concepts and perceptual states. What makes such an individuative link possible is in part the existence of perceptual states with representational content. I suggest that the representational content of action awareness provides a similar resource for the individuation of certain concepts of mental action. Some concepts are individuated in part by the fact that action awareness gives reason to apply these concepts.

One clause in a formulation of the possession condition for the concept *judging that p* should treat the case of first-person application in the present tense. It should state that, in the absence of good reasons for doubt, an apparent action awareness of his judging a given content gives reason for a thinker to accept *I judge that p*. Here, the action awareness in question has a content to the effect that the thinker is himself judging that *p*. When the action awareness is awareness of a judgement, and a thinker self-ascribes in accordance with this possession condition for first-person ascriptions of judgements, his self-ascriptions are sensitive to the event's being a judgement. Quite generally, making a judgement in accordance with one of the clauses of a possession-condition for a concept in the content of the judgement is a way of coming to know the content of the judgement in question.[8] The action-blind subjects considered in the response to Question I could not, incidentally, exercise this concept of judgement in making ascriptions to themselves (if indeed they could possess concepts at all), since they lack action awareness of their judgements.

An account of possession of these mental-action concepts must also have a clause dealing with third-person ascriptions. To understand third-person ascriptions of these concepts is to have tacit knowledge that their correctness requires the subject of the

[7] For further discussion, see *A Study of Concepts* (Cambridge, Mass.: MIT Press, 1992), Chapter 3, and 'Does Perceptual Experience Have a Nonconceptual Content?', *Journal of Philosophy* 98 (2001) 239-64.
[8] A principle I proposed and argued for in *A Study of Concepts* (Cambridge, Mass.: MIT Press, 1992), p.157ff. and in *Being Known* (Oxford: Oxford University Press).

attribution to be in the same state the thinker is in himself, when a first-person attribution is correct. Under this approach, the bridge from first-person ascriptions to third-person ascriptions is built using tacit knowledge involving grasp of an identity relation. This is a structural feature it shares with certain accounts of the bridge from observation-based to non-observation-based applications of an observational concept, and with account of the bridge from first-person ascriptions of other conscious states to other-person ascriptions.[9] Theorists may, however, present other, competing accounts of the bridge, consistently with accepting the first-person clause I have been offering. I will not pursue this further here, since our main focus is on the first-person clause.

Does the first-person clause I have advocated embody a perceptual model of the self-ascription of certain attitudes? It does not. Action awareness is not perceptual awareness; a subject can have action awareness of something without having any perceptual awareness of it. It is no consequence of the present view that when judging in accordance with the relevant possession conditions, one perceives or observes one's judgements or decisions. Nor does the present view postulate intermediaries which would somehow be an obstacle to knowledge of one's own judgements, decisions and other mental actions.

In the case of genuine perception of material objects and events, one would insist that a subject perceives an object or an event itself in a certain way. Far from perception inserting an intermediary that prevents access to the material objects and events themselves, it is perception that makes possible such access to the events and objects themselves. The same is true of action awareness. We should take the grammar at face value. In the bodily case, the subject is aware of his action itself, his clenching his fist, say, and he is aware of it as his clenching his fist. It is as wrong to think of action awareness as some epistemically problematic intermediary preventing access to the events and objects themselves as it is wrong to think of perceptual experience as an epistemically problematic intermediary between subjects and the world.

---

[9] See C. Peacocke, *The Realm of Reason* (Oxford: Oxford University Press, 2004) pp.48-9; '"Another I": Representing Conscious States, Perception and Others', in *Thought, Reference and Experience: Themes from the Philosophy of Gareth Evans*, ed. J. Bermúdez (Oxford: Oxford University Press, 2005); "Justification, Realism and the Past", *Mind* 114 (2005) 639-70; and my book *Truly Understood*, Chapter 4, forthcoming.

Action awareness that one is φ-ing is a factive notion. It implies that one is φ-ing (arguably it also implies that one knows one is φ-ing).[10] As some of the earlier examples show, there is such a state as mere apparent awareness that one is φ-ing, a state whose content can be false. Someone might argue that all we, as agents, ever have is mere apparent awareness that we are φ-ing. This is a form of the argument from illusion in perception, applied here on the side of action.

The argument in the action case is no more sound than its perceptual cousin. When an apparent action awareness that you are φ-ing stands in the right complex of relations to your φ-ing, the apparent action awareness *is* genuine awareness that you are φ-ing. The complex of relations in question is different from those involved in the perceptual case. The relations in question run predominantly from the mind to the world in the action case, rather than the opposite direction of the perceptual case. But the fallacy involved in the argument from illusion is the same in both the perception and the action cases.

Even if the treatment I am offering is not vulnerable to the argument from illusion, it may be thought that it is still open to the objections McDowell has raised against what he calls 'hybrid' accounts of knowledge.[11] As applied to the present subject matter, the complaint would be that on the offered account, there could be a pair of cases in both of which the subject has the apparent action awareness that entitles him to self-ascribe a mental action, yet in one of these cases the self-ascription is true, and the other is false. The objection, to summarize it, is that if this is possible, the self-ascription cannot amount to knowledge in the first case. This is not a paper about general epistemology, so I will not divert the discussion into what would need to be an extended consideration of the status of hybrid theories. The main message of this paper is the role of action awareness in the knowledgeable self-ascription of mental actions. That message can certainly be incorporated into a McDowellian epistemology if one so wishes. That incorporation would proceed by first insisting that in genuine action awareness that one is

[10] I use the notation 'φ-ing' to formulate these generalizations, but this should not be taken to imply that it is only continuing events of which one can have action awareness. One can have an action awareness of something that does not take time, both in the bodily and in the mental domains. Stopping talking can be an action, and the agent can have an action awareness of it. It is not a continuing event. Judging and deciding are also not temporally extended processes, but the subject can have action awareness of them too.

[11] 'Knowledge and the Internal', rep. in his collection *Meaning, Knowledge and Reality* (Cambridge, Mass.: Harvard University Press, 1998).

φ-ing, the subject's mind is embracing the fact that he is φ-ing. The position would then go on to say that the subject, in judging that he is φ-ing is simply taking this factive state at face value, is endorsing its representational content. Action awareness that one is φ-ing would, on this McDowellian incorporation of the point of this paper, play the same epistemic role in relation to certain self-ascriptions of actions as perceptual awareness that *p* plays, on his account, in attaining perceptual knowledge that *p*. On the McDowellian approach, in the case in which subject has a mere apparent action awareness, the kind of state which gives his reason for making his self-ascription of φ-ing is not the same as the kind in which it is genuine awareness of his φ-ing. So the alleged objections to hybrid theories would not get a grip. I am not endorsing this McDowellian approach. The issues involved in assessing it are orthogonal to the main thesis of this paper. My point is just that the idea that action awareness of our mental events is important for the epistemology of some mental self-ascriptions can be acknowledged on both McDowellian and non-McDowellian positions in general epistemology.

Although action awareness is distinct from perceptual awareness, there is a significant parallelism of abstract structure in the perception and action cases on the view I am advocating. There is a structure of rational entitlement in which the entitling state has representational content; and one can be mistaken about whether the content of the entitling state is correct (or whether it is really a factive state that one is in – the parallelism is equally present on a McDowellian treatment). There is a danger here that we may endorse the following fallacious argument:

> Mental actions are not given to their subject under a perceptual mode of presentation.
>
> Hence,
>
> One possible source of error is absent for mental actions that is present for perceptual beliefs about the external world; that is, self-ascriptions of mental actions have a certain domain of infallibility that perceptual beliefs do not.

The premise of the argument just displayed is true. What follows the 'that is,' in the conclusion is false. An apparent action awareness can have a false content, just as an apparently perceptual experience can have a false content. (In the case of mental action, this is the ground of the possibility of one form of self-deception: it may seem to one that one is forming a belief when in fact one is not.) The fact that action awareness is not perceptual-awareness does not give it any kind of infallibility, however limited, that perceptual awareness lacks. The premise of the fallacious argument rightly alludes to the distinction between action awareness and perceptual awareness. This difference in kind does not by itself produce any kind of philosophically significant restriction on fallibility. If there are restrictions, their sources lie elsewhere.[12]

The modest amount I have said so far about mental actions and concepts of them fits a broadly rationalist model of entitlement. The possession-condition for concepts of mental actions contains a clause about first-person present-tense ascription that says that the thinker has reason for making such ascriptions in the presence of suitable apparent action awareness. This accords with a general model under which an entitlement to make a transition to a given judgement always has some a priori component that is founded in the nature of the contents involved in the judgement and the reasons for it, and in the nature of the mental states involved in the transition. Here the relevant a priori component is found in a transition (it is a form of the relatively a priori). A thinker is entitled to take the content of an event of apparent action awareness at face value, in the absence of reasons for doubt. The claim of the existence of some a priori component in every entitlement was the general position I defended in the early chapters of *The Realm of Reason*.

We need, however, to have a much better understanding of how exactly apparent action awareness provides a thinker with entitlement to make judgements about his own actions. The understanding we seek should explain how relying on apparent awareness furthers the goal of making judgements that are true.

[12] I may have been guilty of the fallacy identified in this paragraph. There is a whiff of it in my contribution to a symposium with Tyler Burge on self-knowledge. See my 'Entitlement, Self-Knowledge and Conceptual Redeployment', *Proceedings of the Aristotelian Society* 96 (1996) 117-58, esp. p. 126: "…brute error is impossible. It is impossible precisely because, in these psychological self-ascriptions, there is nothing that plays the role that experience plays in genuine observational knowledge of physical objects."

In the second chapter of *The Realm of Reason*, I distinguished three levels at which one can characterize the entitlement relation. There is, first, the level of instances of the relation. There is, next, a second level of true generalizations about the relation, generalizations that have as instances truths at the first level of characterization. At a third level are principles which explain why those generalizations at the second level are true (and thereby also explain the instances). The third level, as the explanatory level, is the one we should seek to elaborate further in the case of action awareness and the self-ascription of mental actions.

What makes an apparent action awareness one of clenching one's fist, or raising one's arm, or judging or deciding some particular thing, is that, when these and the subject's other mental states are properly connected to the world, they are caused by events (tryings) that cause a clenching of the first, a raising of one's arm, or a judging or deciding of some particular content. That is, the mental states of apparent action awareness are relationally, and in a certain sense externally, individuated. What makes them the states they are is the fact that when all is functioning properly, and the states are properly embedded in relation to the subject's other mental states, his body and the external world, they have a cause which also causes what they are as of – what they represent as being correct. My own view is that the easiest way for such complex, relationally individuated states to occur is for states of their kind to have evolved by a selection process, one which favours the occurrence of those states whose representational content is correct. In taking apparent action awareness at face value, one is judging that things have come about in what is in fact the easiest way for them to come about.

Under this approach, once again it appears that although action awareness is distinct from perceptual awareness, the structure and underlying explanation of entitlement relations involved in relying on action awareness is arguably the same as that underlying perceptual entitlement. The outline just given of why there is an entitlement to take certain action awarenesses at face value is entirely parallel to an argument that there is an entitlement to take certain observational contents of apparent perceptual experience at face value.

This outline of how action awareness entitles a thinker to make self-ascriptions of bodily and mental actions is given for the neo-rationalist approach to entitlement that I myself favour. That approach is opposed to purely reliabilist accounts of entitlement that do not include rationality requirements that are distinct from considerations of reliability. But I should note that pure reliabilists, and no doubt reliabilists of other stripes, could equally accept the importance of taking apparent action awareness at face value in the account of how we come to know our own mental actions. Taking apparent action awareness at face value is not at all something proprietary to neo-rationalists; it can serve many other comers too.

III Is the action awareness model open to the same objections as perceptual models of introspection and awareness?

It is widely held amongst current philosophers of mind that models of introspection that treat it as a form of perception are untenable. I have repeatedly emphasized that action awareness is not perceptual awareness. But action awareness, as a source of self-knowledge, does involve a conscious state that stands in complex causal relations to what it is an awareness of. Action awareness is also, as I have equally emphasized, to be sharply distinguished from judgement that one is performing a certain action; and is also to be distinguished from awareness merely of trying to perform the action. So there is a pressing question: do the objections to perceptual models of introspection, suitably adapted, apply equally to action awareness models of first-person knowledge of mental action?

One of the most interesting and general arguments against perceptual models of introspective knowledge has been developed by Sydney Shoemaker in the second of his Royce Lectures, 'Self-knowledge and "inner sense" ', in the lecture entitled 'The broad perceptual model'.[13] His discussion of the perceptual model of introspection includes the following theses:

[13]Reprinted in his collection *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1996), Chapter 11.

Shoemaker's Thesis (1): Under the perceptual model, "the existence of these [perceptually known] states and events is independent of their being known in this way, and even of there existing the mechanisms that make such knowledge possible" (224-5).

Shoemaker calls his Thesis (1) 'the Independence Condition' (my capitals). I agree that the Independence Condition must be a commitment of any conception worthy of being called perceptual.

Shoemaker's Thesis (2): The Independence Condition implies the possibility of what Shoemaker calls 'self-blindness'.

To be self-blind with respect to certain mental facts or phenomena is to be able to conceive of them – "just as the person who is literally blind will be able to conceive of those states of affairs she is unable to learn about visually" (226) - but not to have introspective access to them. The possibility of such self-blindness, Shoemaker writes, "I take to be a consequence of the independence condition that is built into the broad perceptual model of self-knowledge" (226). I call his Thesis (2) "the Thesis of the Independence/Self-Blindness Link", or "the Link Thesis" for short.

Shoemaker's Thesis (3): Self-blindness is not a genuine possibility in respect of pains; nor in respect of perceptual experience; nor in respect of the will and intentional action; nor in respect of beliefs (sections II, III, IV and V of the Second Lecture respectively).

He elaborates: "…it is of the essence of many kinds of mental states and phenomena to reveal themselves to introspection…" (242). It follows from Shoemaker's Thesis (3), together with Thesis (2), that the Independence Condition is false for pains, perceptual experience, the will, intentional action and beliefs. It also follows in turn by modus tollens from Thesis (1) that the perceptual model of introspection is false.

Shoemaker's Thesis (4): The correct account of the relation of these mental events and states (pain, experiences, intentions, actions, beliefs) to awareness of them needs to draw on the distinction between the core realization of a state and its total realization (242-3).

The core realization comes and goes as the mental state comes and goes. "The total realization will be the core realization plus those relatively permanent features of the organism, features of the way its brain is "wired", which enable the core realization to play [the causal role associated with that state]" (242-3). Adding rationality, intelligence, and possession of the concept of the concept of belief to a first-order belief enables the core realization of the first-order belief to play a more encompassing role. When this surrounding material is present, a first-order belief and the second-order belief that one has that belief have the same core realization. The total realization of the first-order belief is a proper part of the total realization of the self-ascriptive belief that one has the first-order belief (243).

If Shoemaker's arguments in his Theses (1) through (4) are sound, their applicability is not restricted to the perceptual model of introspection. They apply to any subject-matter for which the Independence Condition is fulfilled, and for which self-blindness is not a possibility. This generalizability of Shoemaker's argument is part of its interest and challenge.

It certainly appears that, if the argument is sound, it must generalize to apply against the action awareness account of our knowledge of our own actions (bodily or mental). Action awareness of a particular action is certainly distinct from the action itself. The real or apparent action awareness lies on a different causal pathway from the action itself. The awareness is caused by an initial trying, or some initiating event, which trying or event also causes the effects (the arm's rising) that are required for there to be an action of the kind in question. Even if there is an argument that tryings must, at least in central cases, involve awareness of those tryings, the trying and the awareness of trying is distinct from action awareness. The relation between some constitutive components of the action and the action awareness of the action is causal. It does not seem to be an

option to say that there are no causal-explanatory elements at all in the action awareness account.

But this then seems to leave it at least metaphysically possible that there be actions without the distinctive kind of action awareness that we enjoy. This is precisely the case of action-blindness we considered in addressing Question I above. What makes an event in that envisaged world an action is the fulfillment of the same condition as makes something an action in the actual world: it is caused in the right kind of way by a trying. The actions in this non-actual world would be explained by their agents' contentful intentional states (conceptual or nonconceptual). To fail to acknowledge a category of actions in this possible world would be to miss an explanatorily significant category of events. Action-blind subjects would have to know about even their own bodily and mental actions in ways in ways in which they learn about other events. Even when the formation of one of their beliefs is explained by their other mental states, in wholly intelligible ways, they have no distinctive awareness that they have formed that belief. These subjects would in some respects be as opaque to themselves as another person may be to them. Still, it should not be denied that some bodily events and some mental events in this imagined world are actions, and are appropriately explained by the subject's mental states. There is a plausible case to be made that there can be actions without action awareness, and if this is right, then Shoemaker's Independence Condition is met for actions.

So if Shoemaker's argument is sound, it would follow that the action awareness account is committed to the possibility of self-blindness in respect of such mental actions as judgements, decisions, and the rest. That is what his Thesis (2), the Thesis of the Independence/Self-Blindness Link implies. It is this Link Thesis on which we need to focus in assessing the bearing of Shoemaker's argument on the action awareness account of knowledge of our mental actions.

Whenever something is impossible, one should ask: what is the explanation of the impossibility? If self-blindness is not possible in respect of certain states and events, it may be that the explanation of the impossibility traces to the conditions required for possessing concepts of those states and events, rather than being explained by the failure of the Independence Condition. Actually it seems to me that further reflection on

Shoemaker's own initial illustration of a genuine case of self-blindness, of the genuinely blind person who is able to conceive of the states of affairs that she cannot see to obtain, supports this alternative explanation. The blind person can conceive of objective states of affairs involving objects, events, their properties and spatial relations only because she is capable of perceiving these things and properties in at least some other sense modality – by touch and hearing, for instance (or else because she was once able to see, and knows what it would be to have visual experience of objective states of affairs). If we are asked to entertain the possibility of someone who is supposed to have the conception of material, spatial objects and events whilst also lacking all such perceptual faculties, and lacking all knowledge of what it would be like to have them, it seems reasonable to question whether this is a genuine possibility. It is such faculties that make possible the thinker's possession of concepts of objects and events that may be perceived in one or more sense modalities. If this is so, then there could not be someone who is capable of no perceptual states at all, yet has the concept of objects and events he cannot perceive. The explanation of this impossibility has, however, nothing to do with failure of the Independence Condition. The Independence Condition holds as strongly as ever for conditions concerning external objects, events, and many of their properties and relations. It would be quite wrong to move from the impossibility of someone who both lacks all perceptual faculties and who conceives of objects and events he cannot perceive to the conclusion that the existence of material objects and events is not independent of our ability to conceive of them, to perceive them, or to know of them. Their existence is so independent, in all these respects.

Structurally, the position here is as follows. The claim of the possibility of self-blindness with respect to some states of affairs is a claim of the form $\Diamond (p \,\&\, \sim q)$: it's possible that the subject has the concept of those states of affairs and yet does not have a certain kind of access to them. When self-blindness is not possible, we have something of the form $\sim\Diamond (p \,\&\, \sim q)$ holding. A proposition of that last form is equivalent to the corresponding proposition of the form $\Box (p \rightarrow q)$. The explanation of this necessity's holding may simply be that, necessarily, whenever the conditions for the subject's possessing concepts of those states of affairs hold, the subject also has a certain kind of access to them. Such access may be involved in the possession conditions for the

concepts in question. This can all be true consistently with the Independence Condition's still holding for the states of affairs in question.

A case which seems to me clearly to exemplify this possibility is that of pains and beliefs about pains. An animal can have real pains (not just some surrogate or proto-pains), without having the concept of pain, and hence without having any ability to think about its pains as pains. The existence of pains is independent of their being known about, as the Independence Condition requires. Self-blindness is nevertheless arguably impossible for the state of being in pain. The explanation for this is the widely-accepted point that part of what is involved in having the concept of pain is a willingness to judge, and judge knowledgeably, that one is in pain when one is in pain, where the pain itself makes rational the thinker's judgement. This explanation does indeed not have anything to do with failure of the Independence Condition.

These points also highlight the fact that the sense in which pain is something essentially open to introspection - a consideration Shoemaker uses in the intuitive defence of his Thesis (3) - is to be distinguished from the claim that its nature is constitutively dependent on what its possessor would judge about it in specified circumstances. Introspection is a matter of the occupation and direction of attention, rather than something to be characterized at the level of judgement.

This consideration of the case of pain shows two things:

(a) There are relatively uncontroversial instances in which we have the Independence Condition holding, consistently with the impossibility of self-blindness. It follows that we cannot take the failure of the Independence Condition as the explanation of an impossibility of self-blindness. Shoemaker writes of introspection, contrasting it with perception, that "the reality known and the faculty for knowing it are, as it were, made for each other – neither could be what it is without the other" (245). We are committed to disagreeing with this in one direction: pain could be what it is independently of the presence of the capacity for, and the nature of, thought about pain. The concept of pain is, however, certainly made for knowing about pains. The explanation of the impossibility of self-blindness in the case of pain has more to do with the nature of the concept *pain* than with the nature of pain. It would be wrong, however, to say that the explanation has nothing to do with the nature of pain itself. It is because

pains are conscious, subjective events that pain itself is capable of featuring in the possession-condition for the concept *pain*.

(b) The second lesson is that if, as is also widely accepted, we do not perceive our pains but simply experience them, the Independence Condition can hold even in a case in which the perceptual model itself fails.

An explanation of the impossibility of self-blindness in the case of one's own mental actions is analogous in some respects to that just given for the case of pain, and is disanalogous in others. The explanation is partially analogous in respect of the role played by the possession conditions for such concepts as those of judgement and decision. To possess the concept of judgement involves applying it to oneself in response to one's action awareness of one's own judgements. If a thinker is capable of doing this, he will not be self-blind in respect of his mental actions. His ability to conceive of judgements, decisions and other mental actions as such is constitutively dependent upon his ability to come to know of them in certain way. (A thinker might lose the ability to have states of action awareness of his mental actions, or actions of given type, just as someone may become blind, or blind to certain types of states of affairs. Provided the thinker still knows what it is to be in such states, the corresponding concepts of the states are still available to him.)

It would be an objection to this account of the nature and limits of the impossibility of self-blindness in the case of mental actions if there were a different account of possession of the concepts of judgement, decision and other mental action-types, an account that does not give an essential, constitutive role to action awareness. I do not know how such an account might run. Could an alternative account talk of the thinker's tacit knowledge of an individuating role for judgement, or decision, or some other action-type, in a psychological economy? Such tacit knowledge seems unnecessary in simply making a knowledgeable present-tense self-ascription of an action in rational response to an action awareness of one's performing such an action. For third-person (or other-tense) ascriptions, once one has a the role of action awareness in the first-person, present tense case, a thinker's understanding of the other cases can consist simply in his tacit knowledge that they are correct if their subject is in the same state as someone who is genuinely action-aware of his performance of the action-type in question. In my

judgement, this description of the tacit knowledge is more faithful to what has to be explained than attribution of tacit knowledge of a quite specific psychological role for the action-type in question. Insofar as ordinary thinkers are able to reach conclusions about the role of a mental action-type in a thinker's psychology, it is by way of application of this identity-condition.

A major respect in which the cases of pain and action awareness are disanalogous is that in making a self-ascription on the basis of action awareness, a subject is endorsing the content of representational state. Pain is not, in my view, a representational state (or at least, it is not necessary for the purposes of this account that it be so, unlike the case of action awareness). This difference means that we need an account that addresses the question of why we are entitled to take the representational content in question at face value, as touched upon in the preceding section. Once again, though action awareness is not perceptual awareness, the need for such an account is something shared with the case in which a perceptual experience is legitimately taken at face value.

To summarize this critique to this point: (a) There are counterexamples to Shoemaker's thesis that the Independence Condition implies the possibility of self-blindness; (b) there are alternative explanations of the impossibility of self-blindness, to the extent that it is impossible, consistently with rejection of the perceptual model of introspection; and (c) the explanation of the impossibility of self-blindness has more to do with the nature of the concepts involved in thought about these mental states and events, than in the nature of the events themselves.

What, however, of Shoemaker's own positive explanation of the impossibility of self-blindness in the cases he discusses? There is some reason to doubt that the distinction between the core and the total realization of a state, and Shoemaker's proposal about its extension in cases of introspective knowledge, can do quite the work he requires of it. Shoemaker's view is that the core realization – the realizing state that comes and goes as what it realizes comes and goes – is the same for the mental state thought about and the self-ascription of the state. But since a judgement that one is in a certain kind of mental state requires employment of one's concept of that state (and of oneself, and of the present), the structured state that realizes this judgement is much more plausibly identified as something causally downstream from the mental state that verifies the

content of the judgement about oneself as correct. The realizing state must have sufficient structure for it to realize judgement of a structured intentional content, involving concepts combined a quite specific way. I will not pursue this further here, partly from limitations of space, and partly because the issues are not specific to issues of awareness and self-knowledge. I just note that this second objection will be compelling to those who see something in the arguments, marshaled some years ago in debates about the language of thought, to the effect that the causal-explanatory powers of states with intentional content require corresponding structure in their realizing states.[14] While it is true that Shoemaker includes in his total realization whatever it is that realizes possession of particular concepts, that would not be enough to meet the concerns of these critics. When someone makes the judgement that he is in pain, it is not merely that he possesses the concept of pain. The state that realizes his judgement must also realize the activation or use of his concept of pain, and thus be ready for inferential interactions involving the concept in other premises. Simply being in pain seems to fall short of that. The distinction becomes vivid when for, instance, one thinks one is in pain when the dentist approaches with some terrifying instrument. The dentist then says "I haven't even touched you yet!". In the patient's rush to judgement, he judges that he's in pain, and the realization of this will involve the activation of concepts, and the placing of symbols for the concepts, suitably combined, into the 'belief-box' on theories endorsing the existence of a language of thought. But the subject is not really in pain (nor is a possession-condition relating possession of the concept of pain to the occurrence of pain undermined by such impulsive cases). The most natural treatment of such examples is to say that, even for core realizations, the core realization of pain is distinct from the core realization of the judgement that one is in pain.

[14] J. Fodor, *The Language of Thought* (New York: Thomas Crowell, 1975); 'Appendix: Why There Still Has to Be a Language of Thought", in J. Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass.: MIT Press, 1987); M. Davies, 'Concepts, Connectionism and the Language of Thought', in *Philosophy and Connectionist Theory*, ed. D. Rumelhart, W. Ramsey and S. Stich (Hillsdale, NJ: Erlbaum, 1991).

IV How does the action awareness account reconcile externalism about intentional content with privileged self-knowledge?

Twenty years have now passed since intensive discussion began about the reconcilability of distinctive self-knowledge with the external character of intentional content. The first decade of such discussions has by no means quieted the objections of those who say that some theories of self-knowledge make such reconciliation impossible, so that we must either abandon those theories, or reject externalism about content. Is the action awareness account of our knowledge of our own mental actions an account that makes such reconciliation impossible? And if it does not, what is its positive account of the nature of the reconciliation?

Doubts about the possibility of reconciliation on certain models of self-knowledge have been concisely articulated, and endorsed, by Crispin Wright, writing about halfway (1996) through this twenty-year period. Wright considers the model of self-knowledge as inner observation, and writes "I want to say that …in the sense in which an image or mental picture can come before the mind, its intentionality cannot".[15] "Both a sunburned arm and a triangle can presented as ordinary objects of observation, and each sustains, *qua* presented under those particular respective concepts, certain internal relations: the sunburned arm to the causes of its being in that condition, and the triangle to, for instance, other particular triangles. And the point is simply that while the identification of the triangle as such can proceed in innocence of its internal relations of the latter kind … recognition of the sunburned arm as just that cannot proceed in like innocence but demands knowledge that its actual causation is as is appropriate to that mode of presentation of it." (Ibid., 343). Wright attributes to Wittgenstein, and finds convincing, the point that "the internal relations to the outer, of whatever sort, are all of the latter - sunburn-style - kind; and hence there is indeed a standing puzzle in the idea that an appropriate characterization of them, incorporating such intentionality, is somehow vouchsafed to their subject by something akin to pure observation" (343).

[15] C. Wright, *Rails to Infinity: Essays on Themes from Wittgenstein's Philosophical Investigations* (Cambridge, Mass.: Harvard University Press, 2001), p. 342.

Does Wright's objection apply equally to the action awareness account? In the perceptual case, as Wright implies, there is a partition between properties such as that of being sunburned that cannot be known to be instantiated simply by taking perception at face value, and those such as shape, colour, orientation, surface texture, and so forth, which can be known to be instantiated simply by taking perceptual experience at face value. (A state's representing these latter properties is also a matter of its external relations, incidentally, a fact that should give us pause about the direction in which the argument is going.) An analogous partition of properties, as thought about in given ways, can be made for action awareness. I cannot, from action awareness alone, come to know that the copying machine whose lid I am closing was manufactured in Taiwan. Knowing that requires knowledge of its history that is not given in action awareness. But action awareness can make available knowledge that I am closing the machine's lid, at a certain speed, with a certain force, and that I am doing it now. So, in the case of mental actions, the crucial question to address is this: is the intentional content of a mental event or state to be grouped with the property of being made in Taiwan, or with the properties which you can know about simply from your action awareness?

The intuitive, pre-theoretical answer to this question is that we have an action awareness of the full intentional content of our judgements, decisions, and other mental actions. We are aware that we are judging that New York is hot in the summer; we are aware that we are deciding to spend the summer in a cooler place. A judgement may also be a manifestation of a neurosis, may be an unconscious excuse for not staying in New York, or many other things that are to be grouped with the machine's being made in Taiwan. But the intentional content of the judgement, decision or whatever mental action-type is in question does seem to be so available. What is the explanation of this fact? Whatever the explanation, it will have to have a certain generality. When we know what we are judging or deciding, on the basis of action awareness, we know the content of our judgement or decision, whatever its conceptual constituents. You can have an action awareness of your judging that *p*, whatever the content *p* may be, whether the conceptual content *p* is observational, theoretical, moral, or anything else.

When you judge, on the basis of an action awareness, that you judge that New York is hot in the summer, you are thinking of yourself as having an attitude that

involves a certain concept, that of being hot. You think of that concept in a particular way. There are many different ways of thinking of a concept (as there are of thinking of anything at all): but in this second-order judgement about your first-order attitude, you are thinking of the concept *hot* under its canonical concept. The canonical concept of a concept F, can(F), is something made available by the concept F itself. There are various ways of elaborating this kind of availability in more detail, the various ways corresponding to various theories of how concepts are individuated. Suppose, as a starting point, we think that a concept is individuated by what has to be tacitly known about the condition for something to fall under that concept. Then we can state what is distinctive about the canonical concept of a concept thus:

(*) For an arbitrary concept C to fall under the canonical concept of the concept F (under can(F)) is for C to be such that: the fundamental condition for something to fall under C is the same as the fundamental condition for an object to fall under F.

We can call this a *leverage* account: it leverages an account of the condition for something to be F into an account of a particular higher-level concept of that concept. I am taking it that a thinker could not have the concept F itself (as opposed to some mode of presentation of F) in the content tacitly known in knowing what (*) states unless he has enough of a grasp of the reference-condition for F to be attributed with attitudes containing F in their content.

Under the leverage approach, the canonical concept of a concept F is unique. Under the model of the tacit knowledge of reference-conditions, there is only one canonical concept of the concept F built according to the pattern of (*). Clearly also only concepts can have such leverage-involving canonical concepts of themselves, since the very condition for something to be in the extension of the canonical concept of F requires it to be a concept. For something x that is not a concept, *the canonical concept of x* is something that is not, along these lines at least, well-defined.

This leverage account has two consequences.

(i) Reasons, in given circumstances, for accepting or rejecting particular contents containing the concept F become reasons for accepting or rejecting corresponding

contents containing the canonical concept of F. If something is a reason for judging Fa, in given circumstances, then it is equally a reason for judging that a falls under the concept thought of as the canonical concept of F. Since the fundamental condition for a concept to fall under the canonical concept of F is that it has the same satisfaction-condition as F, any reason for judging that something is F will equally be a reason for judging that it falls under the concept picked out by the canonical concept of F.

(ii) Any externalist features in the individuation of the concept F will be inherited by the concept *the canonical concept of F*.

We can now focus on the transition from a thinker's having

an action awareness of his judging that New York is hot in the summer

to his judgement of this intentional content, where the senses referred to are thought about under their canonical senses:

<Judge>^<I>^(can(<hot>)^can(<New York>)^can(<in the summer>)).

Here '<A>' denotes the sense expressed by A (I omit other formalities). This transition from the action awareness to the judgement is a priori valid. In any context in which the thinker has a genuine, and not merely apparent, action awareness of judging that New York is hot in the summer, it will also be true that he judges that it's the concept *hot* that he judges New York to fall under in the summer. The same applies to the canonical concept of any other concept F in place of the concept *hot*, however externally or historically individuated the concept F may be. The reason-giving state in this transition is as externally-individuated as the content of the judgement that it rationalizes.

This transition in thought from action awareness to judgement is totally different from the transition, unwarranted without further information, from

a perception of a reddish arm

to a judgement

this arm is sunburned.

This latter transition does, just as Wright says, need further information about the causes of the redness on the arm, in a way the preceding transition does not need any further information for its legitimacy. Unlike the transition to a self-ascription of a judgment based on action awareness, the content of the judgement about sunburn involves external factors, about the cause of the redness, whose presence is not ensured by veridical perception of a reddish arm.

There may be a sense of unease about this reconciliation of externalism and the distinctive knowledge of mental actions, a sense that there is some kind of cheating going on. This sense may stem from the thought that the treatment given here is like that of someone who insists, correctly, that the recognitional concept of water is externally individuated, and that we know our thoughts are water-thoughts. There is a clear sense in which one can possess a recognitional concept of water without knowing which liquid it is, in the sense of not knowing its chemical composition. Does a similar objection apply against the account I have offered of action awareness of the conceptual contents one is judging? Is the account consistent with the thinker's not knowing which concepts are in question?

I reply that because one is employing the canonical concept of a concept in making judgements about the contents of one's thoughts, one does, by contrast with the chemical characterization of water, know which concept is in question. It is precisely the force of the leverage accounts to make it clear that any such seeming gap is really closed. Under the leverage accounts, you know as much about which concept is in question when you think of it as *the concept F* as there is to know. All the conditions that contribute to the individuation of the concept F itself contribute to the individuation of the canonical concept of F. From the leverage accounts, as noted, reasons for making first-order judgements containing the concept F are transmitted to reasons for making suitably corresponding judgements containing the higher-order concept *the concept F*. Any requirements on knowing which concept that are met when one is simply using the

concept F will equally be met when one uses the canonical concept of the concept F. It is always an answer to the query "Which concept is in question?" to say "It's the concept F", where this answer employs the canonical concept of F, rather than some descriptive mode such as "the concept discussed in Chapter 6 of this book".

There will, for any given concept, be empirical conditions met by a given thinker who employs that concept, conditions not extractible simply from the nature of the concept itself. They will include such matters as the nature of that particular thinker's mental representations underlying his possession of the concept, and the particular computational procedures involving it that he employs. There is manifestly an important area of study that consists in the empirical investigation of these empirical matters involving concepts as possessed by particular thinkers. But precisely because these empirical conditions can vary across thinkers that share the same concept, these empirical conditions are not what constitute the nature of the concept itself. A thinker's ignorance of these empirical matters of mental representation does not impugn her knowledge of which concept is in question when she thinks "I judge that New York is hot in the summer".

It is a striking feature of the canonical concept of a concept that it has two characteristics whose coinstantiation rests on a merely empirical truth. The canonical concept has the individuating properties specified in the leverage account. Our minds and conscious states are also such that we can rationally apply the canonical concept of a concept in response to conscious states, such as action awareness and passive thinking, whose content involves the very concept of which it is a canonical concept. It seems to be a precondition of rational, critical thought that these two characteristics go together. Rationality requires us, on occasion, to consider for instance whether our conscious judgement that Fa was made in an epistemically responsible fashion. Investigation of this issue involves drawing on our tacit knowledge (or tacit partial knowledge) of the condition for something to be F. Such a rational exercise of thought is possible only because the canonical concept of F is one we can apply in rational response to conscious mental states whose intentional content contains the concept F itself.

Canonical concepts of concepts are far from the only concepts some of whose distinctive applications rest on empirical facts. The way we think of a type of bodily

movement, when we perceive it made by someone else, yet also perceive it as an action of a type that we ourselves could make, provides another type of example. No doubt the underlying ground of the possibility of such concepts involves the now-famous ‘mirror neurons’ identified by Rizzolati and his colleagues. It is an empirical matter that there are such representations in our psychology. They make possible much that would not otherwise be possible. To deny the existence of ways of coming to apply concepts that rely on empirical facts would rule out large tracts of human thought and experience. This applies equally to our ability to know about the intentional content of our own mental actions and our other conscious states.