

TRANSPARENCY AND PRACTICAL KNOWLEDGE

JOHANNES ROESSLER

[FIRST DRAFT]

Transparency

According to Elizabeth Anscombe, in doing something intentionally an agent is in a position to know what she is doing in a distinctive way: her knowledge is not acquired by observation, rather what is known here is known ‘by being the content of an intention’. A natural reading of this suggestion runs as follows. Consider two kinds of question we can ask about our own future actions. The first is the *practical* question: What am I going to do? or what shall I do? where this is aimed at eliciting an expression of an intention for the future. The second kind of question is what might be called the *factual* question: What will I do? where this is aimed at eliciting a belief about the future. One claim that might be made here is that answering the first question can be an excellent way of answering the second question. Borrowing an expression from Richard Moran’s discussion of self-ascriptions of belief, we might say that the factual question can be *transparent* to the practical question, where this means at least the following: answering the practical question can yield a knowledgeable answer to the factual question. Call this the transparency thesis. This is a claim about a possible source of knowledge of future actions. We might call knowledge made available in this way practical knowledge, in the sense that it has its source in the agent’s answer to a practical question. Now when Anscombe’s says that what is know to agents is known ‘by being the content of an intention’, on first encounter, it may seem that she intends this as an endorsement of the transparency thesis: as a claim about the source of agents’ knowledge.

If we accept the transparency thesis then it seems there is a serious philosophical question we face of the familiar type ‘how is x possible, given y?’, where y seems to present an obstacle to x being the case. Or rather we face two such questions, or sets of questions, one relating to future actions and one relating to current actions. The future action question is: how is it possible for an agent to know what she will in fact do simply by expressing an intention for the future? Of course one kind of reason for finding this puzzling may be general concerns about the possibility of knowing what will happen in the future. But here I will only be concerned with a second kind of apparent obstacle: even if we allow for the possibility in principle of knowledgeable beliefs about future events, one might expect that such knowledge has to be based on evidence. Relative to that assumption, it can seem puzzling how it is possible to acquire such knowledge simply by expressing future intentions. The second ‘how possible’ question is the current action question: how is it possible to know what one is currently doing simply by expressing the intention informing one’s activity, without relying on observation? One way of putting a possible obstacle here might be in terms of a competition between the authority of practical knowledge and that of observational knowledge. Suppose we agree that intention provides the agent with a kind of foresight concerning her future actions. The fact remains, though, that, as Anscombe notes, ‘orders can be disobeyed and intentions fail to get executed’ (82). Now in the case of current or completed bodily actions success typically requires the occurrence of events, such bodily movements and their effects, that can be known by observation. And one might feel that in a direct confrontation between the authority of practical foresight and that of observation, the latter must surely prevail. This in turn may lead one to feel that there would be something cognitively irresponsible, as it were, about forming non-observational beliefs about your

current or just completed activities, when they are, or have been, unfolding right before your eyes.

Now in Anscombe's view, I think, these kinds of questions rest on a mistake — so it would also be a mistake seriously to engage with them. At least, this is one kind of response that can be extracted from Anscombe. According to this response, if you are worried about these sorts of questions you are probably in thrall to the 'incurably contemplative conception of knowledge' dominating modern philosophy. What the questions invite us to do, in effect, is to provide a substantive epistemology of practical knowledge — a philosophical account of how transparency provides a source of knowledgeable beliefs. Once we realize that agents' knowledge is not another mode of contemplative knowledge but practical, in a more technical sense, we will be freed of any requirement to provide a substantive epistemology. In Anscombe's words, we will emerge from the 'utter darkness' in which we found ourselves when trying to fit agent's knowledge into the contemplative mould, looking for a 'different mode of contemplative knowledge in acting, as if there were a very queer and special sort of seeing eye in the middle of acting'. A claim to practical knowledge, in Anscombe's more technical sense, is a claim that is not rendered incorrect or mistaken by one's failure to execute the intention it expresses. Of course, failure does mean that practical knowledge is not available: you cannot be said to know, even in the technical sense, that you are pressing button A when you are pressing button B. But Anscombe insists this does not mean the *claim* to practical knowledge of pressing button A was mistaken. The mistake — the *only* mistake — occurred in the performance, in your failing to do what you said you were doing. What this suggests, crudely put, is that agent's knowledge is to be conceived not as knowledge *derived* from intention but rather as knowledge *constituted* by (successfully

executed) intentions. Practical knowledge, in this more technical sense, does not invite the project of providing a substantive epistemology, for the simple reason that such knowledge does not involve or entail belief. Of course there will be questions to be raised about the explanation of the epistemic status of a claim to practical knowledge. But these issues will turn not on the source of such knowledge but on its effects: on matters such as the non-accidental production of intended effects, rather than on matters such as the non-accidental formation of true beliefs. One way to summarize this reading would be as follows. Anscombe accepts that if the transparency thesis is right, then it must be possible to give a substantive account of the epistemology of agents' knowledge. She then argues by modus tollens: faced with the epistemological project we find ourselves in utter darkness, so we should reject or modify the transparency thesis, in line with her technical account of practical knowledge.

Recently there has been an interesting tradition of work on agents' knowledge that is inspired by Anscombe, yet opts for a modus ponens at this point. The idea is that the transparency thesis is correct and that it is in fact possible to give a philosophical account of how transparency yields knowledgeable beliefs. So there is no need, as it were, to discard the transparency thesis in favour of Anscombe's more technical notion of practical knowledge. This tradition seems interesting to me for several reasons. One is that I suspect it is right: it seems plausible to me that transparency holds the key to understanding the nature and importance of agent's knowledge. Another reason is that the tradition I have in mind makes an interesting, and I think plausible connection between agents' knowledge and agents' control of what they are doing. For a central theme in this tradition is the intuition that understanding agents' knowledge requires understanding agency; roughly: the sense in which agents participate in what they are doing, rather than

merely being the locus where actions take place. My third reason is of a more critical, diagnostic character. Some of the accounts of the kind of control thought to help elucidate agents' knowledge strike me as wildly implausible. So I am interested in the reasons some philosophers are attracted to these kinds of theories, and of course I am interested in whether it is possible to sustain transparency without extravagant theories of control.

Agents' knowledge and control

To set the scene I want to distinguish three kinds of issues, corresponding to three levels of theorizing about agents' knowledge. First there are what we might call ground-floor questions concerning the scope of agent's knowledge. Suppose you see someone accidentally knocking a glass off a table and you instantly reach out to catch the glass. How should we characterize the source of your awareness of this action? Or suppose you are delivering a message to someone. Are the precise words you use, as opposed to the general gist of the message, something of which you have practical, non-observational knowledge — or rather something you learn in the same way as the recipient of the message, by listening to what you are saying? For another kind of example, consider an episode from the life of Jonathan Miller, the theatre director, as reported by Tony Marcel:

As a young man, [Jonathan Miller] came across a crowd who, he learned, were waiting for the Queen to pass in a car. When the car passed, the crowd cheered and waved. (This happened in the 1950s when royalist adulation was still widespread.) Jonathan says that to his surprise and horror, given his self-image, 'I found myself raising my arm and starting to cheer', which he quickly suppressed. (Marcel 2003)

In reflecting on such examples, we will find ourselves thinking about the way in which agents' knowledge is connected to issues concerning control. So a more general theoretical question in the philosophy of action we can ask is: how should we characterize the kind of control that, intuitively, goes together with agents' knowledge? Which in turn leads to a third level of reflection: If control (in the sense identified at the second level) yields knowledge, how does it do so? How does control provide for knowledge?

One might expect that the answers to the first order questions about the three examples are simply manifest from the perspective of the phenomenology of action. And, one might argue, from that perspective surely the correct thing to say is that of course in all three cases the agents do know what they are doing 'from within'. After all, in finding yourself catching the glass, you are not in the position of sufferers from Anarchic Hand Syndrome who experience the apparently goal-directed movements of one their hands from the perspective of a passive observer, as something that is not under their control. Nor are you in any abnormal state of mind impairing ordinary knowledge of what one is doing (say, somnambulism or hypnosis), and the same seems to be true of the other examples.

But in fact there is disagreement even about the phenomenology. David Velleman, from whom I have borrowed the first two examples, seems to take it to be *obvious* that your knowledge of catching the glass is based on observation. 'When you extend your hand reflexively' he writes 'you react before you know it, but then you observe your reaction.' (196) He also claims that in the second example, which precise words you use is something you learn only by hearing yourself say them. (21) And I think Velleman would

probably deny that Jonathan Miller could have had practical knowledge of his arm raising and cheering, not just because things happened too quickly, but also because, in Velleman's view, practical knowledge is linked to something like 'identification' in Harry Frankfurt's sense, something that is conspicuous by its absence in this example.

Velleman's account of the ground-floor question is underpinned by a certain ideal: the ideal of action, for Velleman, is deliberated or at least considered action. (I borrow this notion from Bernard Williams 1995.) Now there may be things to be said for this ideal; but I want to suggest understanding practical knowledge does not require embracing the ideal. It seems to me there are two kinds of fallacy at work in Velleman's description of the examples. (I will focus on the first example.) The first fallacy concerns the temporal relation between action and awareness. Velleman maintains that 'behaviour aimed at arresting [the fall of the glass] is initiated and completed before you know it' (191). Suppose we agree that if it is only after the act that you become aware of acting, your knowledge will not be practical knowledge. But one question here is whether Velleman would allow for the coherence of the following kind of scenario. Suppose that at the very same moment t_1 three things happen. One thing is that you acquire a present-tense intention to do something, say to look at your watch now. The second thing is that you begin to act on that intention, say you begin to turn your head in the direction of the watch. The third thing is that you become aware of what you are doing; knowledge that you are looking at the watch becomes available to you.¹ This kind of spontaneous action seems possible, indeed rather common. What is the temporal relation between

¹ One may insist that there will be a small temporal interval between the acquisition of the intention and the initiation of the action; but the point is academic, and in any case, it does not affect the point that knowledge becomes available simultaneously with the acquisition of the intention. (See Haggard for relevant experimental work.)

‘behaviour’ being ‘initiated’ and your awareness? On the face of it, the two things occur concurrently. Then why should not the same be true in the case of the falling glass? The fact that you have no time, while acting, to *articulate* your knowledge is surely neither here nor there. Nor is it obvious why the stimulus-driven nature of the action should make my description ineligible. The fact that you have no *prior* awareness of what you are going to do does not imply that you lack awareness of what you are doing at the time when ‘behaviour is initiated’.

The second fallacy concerns the notion of control. It might be said that the real issue is whether we should think of your act of catching the glass as an intentional action, and that it is here that the stimulus-driven nature of the act is of major importance. Now it is true that Velleman characterizes the action as unintentional (or at least not ‘fully intentional’) (see 189), and as reflexive. So perhaps the force of his intuition that you lack practical knowledge derives from the intuition that the action is not intentional. Now this latter intuition would deserve closer examination.² I do not find it compelling, and I think some of the things Velleman says to make it seem compelling are questionable; for example, his claim that you only become aware of the action after its being completed. (This would be hard to square with what seems like a genuine possibility, of your aborting the action (perhaps because you have always disliked that particular glass) after automatically beginning to reach for it.) But more importantly, I think the question of intentionality is not central to Velleman’s argument at all. Central to Velleman’s account of practical knowledge is a distinction between what he calls mere activity and full-

² The intuition may be supported by appeal to experimental evidence suggesting that in certain cases of stimulus-driven action, say breaking when someone is running on the street in front of you, sub-personal processing devoted to the initiation of the action commences prior to any conscious experience of the stimulus.

blooded action, where, very roughly, the former is a matter of behaviour being caused by suitable beliefs and desires whereas the latter involves, in addition, a distinctive way in which agents control, participate in, and know their actions. What is not obvious is that Velleman's notion of full-bloodedness lines up with the commonsense psychological notion of intentionality. So I think Velleman need not (and should not) insist that you are catching the glass unintentionally. His central claim is this: 'Intuitively speaking what these behaviours lack is that, while directed *at* various things, they are not directed *by* you.' (191) For behaviour to be directed by the agent, Velleman argues, there has to be activity at two levels: goal-directed behaviour and the activity of 'exercising conscious control'. (192) The intuition we are invited to share here is that the agent comes into the picture only in virtue of the extra activity of exercising conscious control. Never mind whether you are reaching for the glass intentionally. The important point is that your reaching is a case of mere activity, that is, an event caused by some belief and desire — something that can be described and explained in entirely impersonal terms, without so much as touching on the existence of an agent of the action. So I think there is a strong top-down influence, as it were, on Velleman's answers to our ground-floor questions.

There is a theoretical background view according to which the notion of control we need to elucidate agents' knowledge is not that of intentional control but of agents' engaging in the second-order activity of exercising conscious control over goal-directed behaviour.

There may be ways of defending this kind of picture, but I think it has little intuitive appeal, and there is a clear fallacy in the way Velleman seeks to bring out its intuitive appeal. The idea is that mere activity lacks agents' participation. But intentional action, whether or not it involves the exercise of conscious control over the action, certainly involves intentional control of the *effects* of the action. (See Hornsby 2004) And the agent is clearly involved in — participates in — that kind of control. In intentionally arresting

the fall of the glass *you* exercise control over the trajectory of the glass. (And you will rightly claim credit for that.) *Deliberative* control over what to do, then, is not the only mode of agents' participation. It is a fallacy to argue that since agents' knowledge implies agents' participation it requires deliberative control.

The third example raises of course special issues. But it might be said that these issues should be seen to complicate, not to supplant, the basic point that Jonathan Miller intentionally raised his arm to cheer the Queen. (Surely, one might argue, this point is an important part of the reason Miller reacted not just with surprise but with horror.) On the other hand, there is a spectrum of ways in which an agent may be epistemically and emotionally alienated from what she is doing, say under conditions of extreme fatigue, drunkenness, perhaps in the pre-onset phase of schizophrenia, and perhaps, one might add, under conditions of mass hysteria (although I suspect it may be pushing things a bit to classify Miller's context under this heading). I cannot pursue these issues here, but I think there is no reason to expect that Velleman's notion of mere activity, as opposed to full-blooded action will prove illuminating in this context. It is not that in such cases, agent's knowledge and participation are simply absent (so that, considered as mere activity, the cases are actually faultless). Rather both agents' knowledge and participation seem to be impaired or defective, and understanding these various (and different) kinds of alienation requires understanding the source and character of the defect.

Knowledge of future actions

We have found no reason so far to think that the notion of control we need in explaining agents knowledge is anything other than the commonsense-psychological notion of intentional control. Moving up a further level of reflection, then, the question is how

control, in that sense, yields knowledge. I want to focus first on the case of knowledge of future actions. According to the transparency thesis, you can gain knowledge of what you will in fact be doing tomorrow by answering the practical question of what you shall do, or what you are going to do, tomorrow. A striking way of developing this is due to Richard Moran:

‘It is as an expression of the authority of reason (..) that [the rational agent] can and must answer the question of his belief or action by reflection on the reasons in favour of this belief or action. To do otherwise would be for him to take the course of his belief or his intentional action to be up to something other than his sense of the best reasons, and if he thinks *that* then there’s no point in his deliberating about what to do. Indeed there is no point in calling it “deliberation” any more, if he takes it to be an open question whether this activity will determine what he actually does or believes.’ (2001, p. 127)

We can distinguish three claims here. One is the general idea (of broadly Kantian inspiration, as Moran notes) that agents’ conception of their future intentional actions should be seen as a commitment integral to practical thinking, rather than as a conclusion based on theoretical evidence. Understanding the possibility of non-inductive knowledge of future actions, then, requires understanding how practical thinking can provide an adequate basis for factual belief. The second claim is that the practical thinking in question is, specifically, deliberating on what to do by reflecting on one’s practical reasons. The third claim might be put by saying that the practical point of view is theoretically impenetrable. If you decide to write a particular letter tomorrow you incur a

commitment that you will in fact write it, and you cannot consistently embrace this commitment while simultaneously qualifying it on the basis of theoretical evidence.

On this last point, it might be wondered whether Moran does not overstate the case. One can have varying degrees of confidence about one's ability to carry out intended, even deliberated, actions, and one's confidence can be informed quite intelligibly by the evidence available to one, not just concerning possible external obstacles but also concerning internal obstacles such as forgetfulness or weakness of will. (See Bratman) It might be said that if you are not confident you won't forget to write the letter or you won't akratically change your mind about this, then you can't really be *intending* to write it (as opposed, say, to wishing that you will). But that seems not true in the case of external obstacles — you can intend to carry out a difficult task without being certain you will succeed; so it is not obvious why it should hold in the case of internal obstacles. True, lack of confidence, in the case of internal obstacles, creates a certain kind of instability: if you really intend to do the thing, then you should take steps to overcome such obstacles (making lists, forming 'implementation intentions' etc) , which in turn should lead to increased confidence that you will do what you intend to do. Nevertheless, it would surely be an illusion to assume that internal obstacles can always be overcome. And it would seem to be a corresponding theoretical mistake to assume that a deliberator is committed to that illusion. Moran acknowledges that the authority of the practical point of view can be 'hedged', as when I know I am akratic with respect to a certain practical question. But he maintains that this will 'compromise the extent to which I can think of my behaviour as intentional action (..)' (128) Akratic actions can be intentional, though, and it is not clear why a deliberator should have to be blind to the fact that it may well be an intentional action on her own part that will foil her plans.

These are interesting and delicate issues (both theoretically and practically), but they are relatively independent of the basic epistemological problem. Suppose someone insists that if your belief about future action is to have an adequate basis it has to be based on evidence. The simplest way to develop this sort of picture would be to appeal to another kind of transparency: by answering the practical question of what you are going to do you put yourself in a position to answer the psychological question of what you intend to do. Then, armed with a knowledgeable answer to that question, you may reason along the following lines: I intend to write a letter tomorrow; on the whole, people do what they intend to do; so it's likely that I will in fact write a letter tomorrow. I take it Moran's response to this sort of view would be that if this is the way you generally think about your future actions you probably need help. You are in state of alienation or estrangement that threatens the stability of your sense of self. It is here that Moran's second claim, about the importance of deliberation, comes into play. For he thinks that the general demand for evidence is inconsistent with crediting the agent with the capacity for practical deliberation. The key claim here is that the transparency of the factual question "what will I do" to the deliberative question "What do I have most reason to do?" is not just an option but a normative requirement, grounded in the agent's understanding of the very point of deliberation. It is not just that by deciding what to do tomorrow you put yourself in a position to answer the question: what will I do? Rather, in answering the deliberative question you simultaneously commit yourself to a particular answer to the factual question. The basic point here is independent of the issue about theoretical penetrability. Even if we somewhat relax Moran's strictures on the latter point, we can agree that unless you show at least a certain degree of confidence that you will do as intended, it becomes difficult, both for you and for others, to sustain a conception of you

as a subject of agency and deliberation. To answer a practical question by deliberation you need to grasp the point of that deliberation, which is to identify reasons that will be effective in determining what you will be doing. Accordingly, a deliberation-based answer to a practical question essentially carries a commitment to a (however qualified) answer to the factual question. And this suggests you cannot consistently think of yourself as an effective deliberator and accept the general demand for theoretical evidence about your future action.

There is of course much more to be said on whether and how this line of argument might explain the possibility of non-inductive knowledge of future action, but I take it it provides at least a promising starting point. What I want to question is the kind of importance Moran's picture accords to deliberation. In effect, he thinks that to develop the (broadly Kantian) first claim — the idea that agents' beliefs about future actions should be seen as a practical commitment —, we need to buy into a (more specifically Kantian) set of contrasts, such as the contrast between justifying reasons and explanation, or decision and discovery. On this view, it is an essential feature of agents' knowledge that it manifests the 'authority of practical reason'. For Moran, the key to transparency is that 'the agent takes the question of what he does [or will be doing] to be answered by his decision as to what is worth pursuing.' (126)

On the face of it, though, you may be able to tell me what you will be doing tomorrow without making any decision and indeed without engaging in deliberation. You may simply express some of your long-standing intentions for tomorrow, say the intention to go to the opera. Perhaps that intention was originally formed by careful deliberation; but all you may now be able to say is that it seemed a good idea at the time. Again, you may

express habitual intentions for things that don't normally get deliberated at all: you will get up, have breakfast, walk the dog, read the paper, and so on. It is not that you are telling me these things are likely to happen, given that they always do. You may genuinely be expressing an intention or plan. There seems to be a kind of certainty about one's future actions that results from never having deliberated about them. True, if pressed, you will no doubt be able to produce various kind of justifying reasons. But it is not clear that these will be the reasons for which you act, nor that you have to regard them as such. And in any case, it seems possible to express intentions for habitual actions without committing yourself to the existence of justifying reasons at all. It does not seem unintelligible to respond to the demand for justifying reasons by saying: 'this is simply what I do'. (Wittgenstein 1953) Even if you find this an unsatisfactory response, its very intelligibility suggests that agents' knowledge may be less tightly linked to agents' 'sense of their best reasons' than Moran makes out.

I want to suggest that the broadly Kantian idea can and should be developed without giving deliberation the very central role it occupies in Moran's account. (Moran's account, like Velleman's, seems committed to regarding deliberated action as the ideal of action.) According to the transparency claim, answering practical questions about future actions can be a source of factual knowledge of future actions. We should not assume or expect that rational agents be generally able to answer practical questions by producing supporting practical reasons. (We do not make that assumption in the theoretical case either.) What is essential is the ability to *express* intentions. A practical question is understood precisely as a question aimed to elicit the expression of an intention. And it is by expressing intentions, whether in solitary thinking or in communication with others, that agents are aware of their own future actions.

Clearly much depends here on what is involved in expressing intentions. It might be said that expressing an intention is itself an intentional activity, and that this means the agent must be aware of which intention she is aiming to express. So expressing an intention presupposes knowledge: specifically, knowledge of the content of the intention to be expressed. This would suggest that the ability to express one's intentions presupposes, and hence cannot explain, agents' knowledge of their intentions; and surely the source of knowledge of future actions should then be located in the source of agents' knowledge of their intentions, rather than in the ability to express intentions. But it seems to me that this account of expressing intentions is neither obligatory nor attractive. True, expressing an intention is itself typically an intentional activity. That is, it is intentional under some description. But it is not clear that the relevant description requires advance knowledge of the content of your intention. For example, if you tell me about your plans for tomorrow the prior intention informing your utterance may simply be to say what you are going to do tomorrow — as opposed to the prior intention to say 'I will go to the opera'. Suppose you can't immediately remember what you were going to do tomorrow. So you try to recollect. It is not clear that the successful completion of that project has to precede the expression of your intention. What you are looking for is the answer to the question 'what was I going to do tomorrow?'; and your utterance 'going to the opera' may precisely be the event of the right answer coming to mind. It is not clear that the latter has to precede the former. As Bernard Williams has put it, 'in their most primitive form expressions of belief are spontaneous': 'in the first instance and in the simplest cases, we are disposed spontaneously to come out with what we believe'. (2002, p. 75) I suggest the same is true in the case of expressions of intentions.

One worry about this proposal is that it might seem to commit to the view Dennett extracts from, or reads into, E.M. Forster's famous line 'How do I know what I think until I see what I say?' — the view that we know about the content of our attitudes by inference from our linguistic dispositions.³ On this view, you see (or perhaps hear) yourself saying 'I am going to the opera', interpret this as an expression of intention and conclude that this is what you intend to do (and thus probably will do). However, this worry rests on an assumption I earlier suggested we have no reason to accept. This is the assumption, operative e.g. in Velleman's work, that agents' knowledge only pertains to pre-meditated aspects of their actions. On this view, to say that you 'find yourself' saying something should be taken as an indication that your knowledge of what you are saying is observational. (Recall the second example considered above.) But normally, I suggest, a more plausible interpretation of this phrase is that you say something *spontaneously*, where this means: intentionally, but informed by an intention that is no sooner acquired than executed, or whose execution begins with its acquisition. So when you tell me 'I am going to the opera' we need to distinguish two intentions. The prior intentions to let me know about your plans for tomorrow and the spontaneous intention to say 'I am going to the opera'. You do not 'find yourself' telling me about your plans, but you may 'find yourself' spontaneously saying 'I am going to the opera'. Consistent with being spontaneous, this latter intention affords agent's knowledge of what you are saying. (So contra Dennett, you are not in the same position as an external interpreter.) Put in

³ 'We often discover what we think (and hence what we mean) by reflecting on what we find ourselves saying — and not correcting.' (So far I agree. But Dennett goes on to say:) 'So we are, at least on those occasions, in the same boat as our external critics and interpreters, encountering a bit of text and putting the best reading on it that we can find.' (1991, p. 245) (Here I disagree: finding oneself saying something can involve agents' knowledge of what one means: knowledge made available by intentionally (if spontaneously) saying something; and such knowledge will not be available to our 'external critics'.)

Williams' helpful terms, your utterance may be deliberate or pre-meditated 'with respect to whether' but not 'with respect to what'.⁴

The suggestion, then, is that it is the ability spontaneously to express our intentions that provides the source of agents' knowledge both of what we intend to do and of what we will be doing.⁵ How does this bear on the epistemological challenge considered earlier? Well it suggests that Moran may be exaggerating the importance of an agent's conception of herself as an effective deliberator. What makes the demand for theoretical evidence inappropriate may not be so much the agents' grasp of the point of deliberation, as her basic grasp of intention and what it means to express an intention. If you express an intention for tomorrow, but refuse to engage with the factual question of what you will be doing tomorrow in any other way than by theoretical reasoning based on evidence, it becomes difficult, for you and others, to sustain a conception of yourself as expressing an intention and commanding a grasp of what it is to have an intention. A fundamental feature of intentions, as opposed to wishes or mere desires, is that they *control* rather than merely influence intentional actions. One way in which this comes out is in that intentions (unlike wishes or desires) pre-empt (further) deliberation as to whether one should act on them. Of course you may always reconsider what you want to do, but the basic 'inertia' of intentions means that if you intend to phi the question of whether to phi

⁴ Williams says 'voluntary with respect to whether but not with respect to what' — but I think my gloss brings out the point he has in mind.

⁵ One important implication of this is that the difference between overt utterances and inner speech becomes relatively insignificant (except perhaps that in the latter case spontaneity 'with respect to what' is not merely optional but of the essence). This in turn may have consequences for the way we should think about the relation between first and third (or second) person perspectives in this area. Another set of issues that would be worth exploring here is how to understand the limitations of agents' knowledge: specifically, whether appeal to someone's inability (for some reason) to express a certain intention may help to explain her inability to be aware of the intention and the intended action.

has been settled. While this may and typically does give rise to deliberative questions as to *how* to implement that intention, the question of *whether* to implement it has been dealt with. (Bratman 1987) Now suppose you express an intention to phi, yet refuse to accept any commitment (however qualified) that you will in fact be phi-ing, unless presented with suitable evidence. You simply insist on being agnostic on the factual issue. This would suggest you don't think of your answer to the practical question as expressing a state controlling your future. So it becomes difficult, for yourself and others, to take your answer to be expressive of an intention at all. (A charitable interpretation might be that you are vacillating between intending to phi and merely wishing you were going to phi.) It is not that your intention provides *evidence* as to what you will be doing, and that you seem to be strangely oblivious to that evidence. If you express an intention, then in doing so you incur a commitment as to what you will be doing that precedes any reflection on evidence, a commitment grounded in your grasp of the point of what you are currently doing: expressing an intention for the future. Thus you cannot consistently think of yourself as answering a practical question and accept the general demand for theoretical evidence about your future action. Put differently, if you do accept the general demand for theoretical evidence you must be in a state of alienation or estrangement, threatening the stability of your sense of self — in particular of your sense of intentional control over your own future. Of course, as with Moran's account, this is hardly a full response to the epistemological challenge. It merely suggests a way of understanding the fact that in the case of agents' knowledge commonsense epistemology resists the demand for evidence. As such, though, it seems a promising starting point.

Knowledge of current actions

There are a number of problems that arise when we try to bring the transparency thesis to bear on knowledge of current actions. One concerns the descriptions under which agents can be said to be thus aware of what they are doing. There are reasons to think that, as Anscombe puts it, the descriptions under which agents know what they are doing are typically 'at a distance from the details of one's movements'. (1957, p. 53) Yet there is also reason to think that the distinction between basic and non-basic actions corresponds to an epistemological distinction between more or less basic knowledge: your knowledge that you have switched on the computer seems to have a more complex source than your knowledge that you have pushed a certain button. (Peacocke 2003) And it can seem as if these principles pull in conflicting directions, the first towards a relatively rich, the second towards a relatively austere account of the description under which agents are (most immediately) aware of what they are doing. I will not pursue this issue here. I will assume that it can be resolved in a way consistent with the view that the content of agents' knowledge matches the content of her intentions; and that the transparency thesis, therefore, is not obviously at odds with the facts. Appeal to transparency certainly promises a parsimonious account of agents' knowledge: the source of such knowledge is provided by the ability to express the intention informing what one is doing; and there is no difference in principle between the case of 'practical foresight' and the case of awareness of actions currently unfolding or just completed.

The problem I want to consider is the following. Perhaps not all of the things we do involve the occurrence of events, let alone bodily movements. (Hornsby 2004) Leaving the phone unanswered is something you may intentionally do, and it is not clear that there is an event that can be identified as your action. Nevertheless, typically intentional action does involve the occurrence of events, and in many cases, minimally successful action

requires the occurrence of bodily movements. The kinds of events relevant here, it might be argued, can be experienced: certainly bodily movements can be objects of perceptual experience, and perhaps mental events such as deliberating about something can be experienced in some sense too. The general challenge facing the transparency claim is that relying on transparency, when experience is available, can look like cognitive recklessness. It can look like relying on evidence concerning the presence of a pig, such as foot prints, buckets of pig-food etc, when the pig is there right in front of you and plainly in view. (see Austen 1962) True, it is not by providing you with evidence that your intention to phi makes it reasonable for you to believe you will phi. The point of the comparison is that in both cases the basis of your belief seems indirect, compared with the kind of direct presence afforded by experience. Intention may provide you with reliable 'practical foresight' of what you will be doing. Nevertheless, having an intention amounts to something less than being presented with the event constituting its successful execution. Thus in relation to current actions, it seems that the authority of intention is surely trumped by the authority of experience. Given that the latter is available to the agent, it would be irresponsible to rely solely on transparency.

The version of the challenge I want to focus on is concerned specifically with perceptual experience: proprioceptive or visual experience of one's bodily movements and their effects on objects in one's environment. To be protected against the charge of cognitive irresponsibility, the argument goes, judgements about one's current bodily actions have to utilize observation; so agents' knowledge cannot be non-observational. An obvious move here would be to try to accommodate the intuition that the availability of experience imposes responsibilities on the agent, without compromising the independent authority of transparency. It might be said that you are entitled to rely on transparency so long as

experience does not present you with evidence to the *contrary* — evidence that you are not succeeding in acting as intended. All that would be required to meet the demands of cognitive responsibility would be to be responsive to evidence of failure, not to rely on evidence of success. But while this may be intuitively attractive, it is clearly not enough to meet the epistemological challenge. It does not begin to address the issue of the relative authority of knowledge made available by intention and knowledge made available by perceptual presence.

We can distinguish three types of response to this challenge, all of them familiar from the case of external world scepticism (there does seem to be a certain affinity or analogy, at a high level of abstraction, between the current challenge and standard sceptical challenges to perceptual knowledge of the external world).

- (1) *Retreat*: One might concede that transparency on its own is insufficient to yield knowledge of current bodily actions and retreat to the view that transparency yields knowledge of what one is intending to do. Knowledge of current bodily actions, it might be said, is to be acquired on the basis of multiple sources: transparency, perceptual experience (perhaps other kinds of experience too).
- (2) *Meet the challenge*: One might try to address the issue by offering a substantive account of the epistemic justification of beliefs based on transparency, an account that enables us to see how transparency can be an autonomous, non-observational source of epistemically justified beliefs as to what one is doing (at least in the absence of contrary perceptual evidence). Thus, Kevin Falvey argues that, for example, when setting about brushing my teeth I have a *right* to assume that the perceptual-motor routines making up the program ‘brushing my teeth’ are running

properly, and that this right in turn ‘gives me the right to employ the expression of intention as a description of what I am doing.’ (2000, 34-5) If this complex of entitlements can be made good, transparency can be seen to be a source of knowledge. Correlatively, reliance on transparency in making judgements about bodily actions will be vindicated.

(3) *Reject the challenge*: In a spirit familiar from Moore’s response to external world scepticism, one might reject the terms of the challenge, by arguing as follows: ‘I know that I am raising my hand [said whilst raising one’s hand], and this knowledge is ‘knowledge in intention’, not based on observation. Now knowing that p is inconsistent with being (seriously) cognitively irresponsible in accepting that p. There can have been no serious cognitive irresponsibility, therefore, in my acquiring the non-observational belief that I am raising my hand.’

It might be wondered whether (2) and (3) are incompatible. They certainly agree that we should reject the sceptic’s central assumption, that the authority of transparency compares unfavourably with the authority of perceptual experience. The basic difference is that (3) simply relies on transparency, whereas (2) aims to explain and vindicate it. According to (2), when you intentionally raise your hand, there are good epistemic reasons for relying on transparency, and it is the existence of these reasons which explain how your reliance on transparency is cognitively responsible. Of course, it is not that you are expected to be able to produce these reasons. We are talking about a philosophical theory of justification or entitlement. You can enjoy an entitlement of the kind identified and defended by such a theory, without being able to appeal to it yourself. What matters, however, is that it would be correct for you to appeal to it, and that this enables us to meet or refute the sceptical challenge. It is here that the Moorean response disagrees. According to (3), the

correct way to deal with the sceptic's assumption is not to develop a substantive account of the epistemic justification of agents' beliefs but to expose its inconsistency with the reality of agents' knowledge.

As it stands, (3) is under-theorized, to put it mildly. So, of course, is (2), but here at least we seem to have some general idea about the kind of story that might be told. What needs to be added to (3) to make it look like a serious option is some theoretical motivation for refusing to engage with the sceptical challenge on its own terms. One such motivation might be a certain kind of naturalism. The issue raised by the sceptical challenge, it might be said, is an idle one: our reliance on transparency is a basic fact about us, and it would be mere pretence to suppose that we do, or could, rely on transparency for the reason that we have a right or entitlement to do so. The kinds of consideration philosophers might adduce in rebutting the sceptical challenge are not, and cannot become, *our* reasons for believing what we do about our own intentional actions. (see Strawson 1985) If this is to be convincing, more would need to be said about the sense in which our reliance on transparency is a 'basic' fact about us, not in need of vindication by a substantive account of epistemic justification. I want to end by presenting, in rough outline, one possible line of argument on this issue that seems at least worthy of further consideration. Needless to say, a sustained discussion of the options is beyond the scope of this paper.

The argument turns on the reasons that might be given for resisting (1). It is of course common ground between (2) and (3) that (1) should be resisted, but it is not clear that they do so for the same reasons. One thing that might be said is wrong with (1) is simply that as a matter of fact agents' beliefs about their actions tend not to be formed in the complex way (1) recommends. As far as our actual beliefs are concerned, then, (1) would

threaten to have unwelcome sceptical consequences. A somewhat stronger reason for rejecting (1) would be scepticism as to whether the combination of awareness of intention and observation of bodily movements could in principle provide an adequate basis for knowledge of intentionally moving one's body. But the strongest response to (1) would be this. What (1) assumes is that from the agent's perspective, intending to move her hand leaves it an open question whether she is in fact moving it. It can be rational for the agent to intend to move her hand but be agnostic on whether the hand is in fact moving. For her knowledge that she is moving the hand is supposed to be result of her independent, observation-based knowledge of the movement. Furthermore, this is so, according to (1), as a completely general matter: not just in special circumstances, say, where you have reason to suspect your hand might be plegic, but in relation to your whole body, all the time. One response to this would be to adapt what I described as Moran's response to someone who is generally agnostic (in the absence of theoretical evidence) about whether he will act as intended: if this is the way you generally think about your bodily movements you probably need help. If in your judgement as to how you are moving your body you insist on relying on theoretical evidence, it becomes difficult, both for you and for others, to sustain a conception of you as a subject of intention, with a grasp of what it is to have and express intentions. For intentions are states that aim to control rather than merely influence one's actions, and they are informed by beliefs about one's practical abilities. If you are not committed to thinking you are moving your hand, it becomes difficult to credit you with the ability to express your intention to move it. I think there is a further parallel with the case of agnosticism about acting in the future as now intended. I suggested that this latter condition would be a state of alienation that threatens the stability of your sense of self: it would undermine your sense of intentional control of your future (which in turn might undermine your very

sense of *having* a future). General reliance on observation of movements in one's judgements about intentional bodily movements might be said to pose a similar threat: it would undermine your sense of controlling your body, which in turn might undermine your awareness of your body *as yours*.⁶

Suppose this very strong line of response to (1) can be properly developed and defended. What I want to suggest is that this might provide a line of argument in support of (3) and, connectedly, an objection to (2). The idea would be that (2) is really in no better position than (1) to appreciate the *practical* source of agents' beliefs and hence the distinctive nature of the authority of agents' knowledge. True, (2) does not demand that agents rely on anything other than transparency itself in acquiring agents' knowledge. Still, according to (2), the authority of this knowledge is underpinned by the existence of good epistemic reasons — a complex of epistemic justifications or entitlements that could be appealed to in rebutting the sceptical challenge. And this is to assume that these are considerations that *could be* an agent's reasons for believing what she does about her intentional actions, an assumption that the strong response to (1) would seem to call into question. If this is right, then the issue raised by the sceptical challenge would seem to be an 'idle' one not just in the sense that its resolution would as a matter of fact have little impact on our belief-forming practices. Rather, it might be argued, we are committed, in virtue of the demands of intentional agency (and perhaps self-awareness), to regarding it as idle.

⁶ Perhaps an illustrative example bearing on this is provided by the tendency of some stroke patients with hemiplegia to disown their plegic limb. See Marcel 2003.