

KNOWING YOUR OWN MIND¹

David Owens *University of Sheffield, United Kingdom*

What is it to “know your own mind”? In ordinary English, this phrase connotes clear headed decisiveness and a firm resolve but in the language of contemporary philosophy, the indecisive and the susceptible can know their own minds just as well as anybody else. In the philosopher’s usage, “knowing your own mind” is just a matter of being able to produce a knowledgeable description of your mental state, whether it be a state of indecision, susceptibility or even confusion. What exercises philosophers is the fact that people seem to produce these descriptions of their own mental lives without any pretence of considering evidence or reasons of any kind and yet these descriptions are treated by the rest of us as authoritative, at least in a wide range of cases. How can this be?

Most of the philosophers exercised by this problem would regard the English phrase “knows his own mind” and its connotations as a mere distraction, as the product of a theoretically unhelpful ambiguity. But to some, the English phrase suggests a fruitful approach to the philosophical problem of self-knowledge. In the twentieth century, Wittgenstein, Sartre and Austin all explored the idea that such mental phenomena as thinking, reasoning, deliberating are, in an important sense, activities which culminate in deeds, in the making of cognitive commitments. Moran sets out to refurbish this tradition, to revive the notion that self-knowledge is special because it is a matter of actively making up your own mind, rather than of passively apprehending it.

Moran’s approach looks most promising when applied to the cognitive parts of our mental lives: our beliefs, judgements, the more intellectual of our desires and emotions. Yet the contemporary discussion of self-knowledge focuses at least as much on experiential states like pain and visual sensations. Moran says at the outset that he thinks these questions about experience require a quite different treatment (p. xxxiii) and focuses his attention on the cognitive. Some of his opponents might regard this as an admission of defeat but the range of phenomena Moran does cover is

¹ Richard Moran, *Authority and Estrangement* (Princeton NJ: Princeton University Press, 2001), xc +

sufficiently impressive to make his view worthy of very serious consideration. In this article, I shan't try to discuss Moran's account of belief, desire and emotion. Instead I will concentrate on his view of our knowledge of our own intentions and of what are doing (or shall do) to execute them.

The Epistemology of Agency

People in a normal frame of mind usually know what they are doing and they know this without having to observe their own behaviour. In Anscombe's example, which Moran adopts, someone pumping water may know that they are operating a pump without observation but not know that they are creating a clicking noise except by hearing the noise (pp. 124-7). In Anscombe's view, this is because making a noise is no part of what they are trying to do. Anscombe acknowledges that one can be wrong in thinking that one is operating a pump (the pump one means to operate might be an hallucination); her point is that our knowledge of our own actions is groundless, not that it is infallible. We can (and often do) know what we are doing without the evidential backing needed for knowledge of someone else's pumping.

As both Anscombe and Moran observe, this epistemic authority extends to our own future agency (p. 88). A person in a normal frame of mind can come to know that they will go to London tomorrow simply by deciding to go to London tomorrow. Again this way of knowing the future is hardly foolproof: an unexpected rail strike or weakness of will may intervene, and empirical evidence as to the reliability of the railways etc. is needed before the decision can be taken. Anscombe's point is that a normal person can know that they will go to London in this way without amassing inductive evidence of their own resoluteness, of their propensity to stick to and execute their own decisions, whereas if they were asked whether someone else who has decided to go to London will actually go, they would need evidence not only about the state of the railways but also about how resolute the person in question is. For Anscombe, the fact that we (directly) *create* this action gives us a special, evidence-independent way of knowing of it. We lack groundless knowledge of what

other people are doing or will do precisely because we don't (directly) create their actions, rather we apprehend them as objects in the empirical world.

In expounding these points, I confined my attention to "people in a normal frame of mind". An example devised by David Velleman shows why this qualification is needed.² I have agreed to meet an old friend in order to patch up a quarrel. As the meeting unfolds, I start to become petulant, raise my voice, provoke my friend into saying some unconscionable things and we part in anger. I later infer that I must have decided, before the meeting, to end the friendship, an objective which I skilfully attained. But at no stage either prior to or during the meeting would I have announced such an intention and even afterwards I unearth it only in an effort to explain what happened. I may end up convinced that this was my intention but only *qua* observer of myself.

I take it Moran has this sort of case in mind when he says that:

Even within a psychoanalytic explanation, it will normally be the case that the contrary thoughts and attitudes which explain the subject's blocked *awareness* of the intention will themselves be reasons for ambivalence *in* his overall intention; that is the intention itself will not be a wholehearted one. Ignorance in such a case will not be mere ignorance, not only because it will be irresistible to look for a *motivation* of sorts to explain it, but because the motivation we then impute to the person must qualify the ascription of the original intention (as conflicted or partial). (p. 57)

Applying this to Velleman's example, I don't fully intend to end the friendship because I can't, to use Moran's favoured expression, *avow* this intention and I can't avow this intention because I am aware of good reasons for not ending the friendship. In general, Moran thinks that a subject is an epistemic authority about the content of his intentions and intentional actions only where the subject's intentions seem well grounded to the subject himself. To the extent that an intention seems to him misguided, the subject's privileged epistemic access to it will be compromised, as will

his sense of being in control of behaviour guided by this intention. To put it another way, he will start to wonder whether this intention is, in the full sense, *his*.³

Moran's epistemology of agency takes off from the fact that we create our actions, we don't just observe them. But why, someone may ask, does the mere fact that we create our actions make us an authority about them? I am unsure whether Moran offers to answer to this question. He might regard it as a brute fact that everyone is an authority about the character of their own creations (including their actions). Alternatively, he might be offering us an explanation for this fact, an explanation which goes as follows: when we act, we act for a reason and the character of our action is determined by the character of our reasons. Subjects (in a normal frame of mind) know the character of their reasons – they know what practical considerations they find convincing and persuasive - and so they know the character of their action. A subject is an authority about what they are doing just in so far as they are an authority about their reasons.

If Moran is telling us the latter story then he must be operating with a highly rationalistic notion of agency. A moderate rationalist would require that truly intentional action be motivated by *a* reason, by the agent's awareness of *some* respect in which the outcome intended would be desirable, without requiring that the agent act on (what he would think of as) his strongest or most powerful reasons. But this moderate rationalism does not suggest a substantial epistemology of agency: we can hardly *explain* how an agent knows what they are doing by supposing that they know what reasons they are *acting on*. To avoid this difficulty we are pushed towards a more extreme rationalism which insists that truly intentional agency is agency determined by the strongest reasons which the agent knows they *have*. Since knowing what reasons you have is not synonymous with knowing what you are doing, there is a substantial epistemology of agency on offer here. Yet such hyper-rationalism runs up against the fact that people intentionally do things they know they have sufficient reason not to do.

² D. Velleman – *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000) pp. 126-7.

³ See also Moran discussing of emotions which one can't avow on p. 93.

A hyper-rationalist can acknowledge that various forms of behaviour approximate to fully intentional and responsible agency but, for him, agency occurs in its pure form only when it accurately reflects the subject's reasons. I am unsure whether Moran is invoking this rationalist notion of agency because I am unsure whether he means to offer us a substantial epistemology of agency. This uncertainty is connected with another interpretative issue: how should we understand Moran's discussion of practical irrationality and, in particular, his notion of akrasia?

Practical Irrationality

On the usual understanding of these terms, akratic or incontinent action is action which is performed even though the agent judges that it would be better for him to be doing something else. Akrasia is a very familiar phenomenon. If I eat more biscuits than I know I'll be able to digest comfortably or remain slumped in front of the TV when I realise I ought to be making lunch, my eating and my lounging are things I do, things which I do deliberately, freely and intentionally, things for which I am fully responsible. In these respects, akratic action is no different from fully continent agency: it is agency *par excellence*.

Understanding akrasia in this way, I find some of the things Moran says about it rather puzzling. For example

when I know that I am akratic with respect to the question before me, that compromises the extent to which I can think of my behaviour as intentional action ... Nor does a person speak with first-person authority about such conditions. (pp. 127-8)

Moran also says that the akratic does not "identify" with their action and that his knowledge of it is empirical and not "ordinary self-knowledge" (p. 67).⁴ The implication is that while the akratic may be able to predict that he is likely to behave akratically and then exercise a sort of self-control by putting obstacles in the way of akratic behaviour (e.g. placing a time-lock on the drinks cabinet), he will lack that

first-person knowledge of and control over his behaviour which the continent possess.⁵

Moran's description fits Velleman's psychoanalytic example very well but as a commentary on everyday akrasia, it is a bit overstated. My partiality to lunch time TV hardly qualifies as an obsession or a compulsion even though it tempts me to do what I know I should not do. I am not overwhelmed by the desire to watch, I choose to indulge it. If I am honest, I won't deny that it is *me* who decides to remain on the sofa. Furthermore, I know perfectly well *why* I take this decision: the lunchtime soap opera is genuinely entertaining and diverting (though, even in my own eyes, other considerations are more pressing). There is no failure of first-person access either to the motivation for my akratic behaviour or to what it gets me to do.

Reflection on such everyday cases of akrasia puts pressure on Moran's notion of an avowal. Moran says that there are two elements to avowal: first, an authoritative awareness of the state of mind avowed and secondly an endorsement of it (pp. 91-2). Consider a Catholic woman, described by Jackson, who has become pregnant after rape.⁶ She judges that she ought not to have an abortion but akratically decides to have an abortion nevertheless and makes plans to attend the abortion clinic next week. Can she avow this intention? She can do more than *report* this intention in the way she might report a third parties' intention but she can't go as far as to *endorse* this intention. What she can do is to *affirm* that she is set on having an abortion, that she has resolved to have an abortion, that she is committed to having an abortion. Moran's notion of an avowal seems to include the idea of endorsement (p. 67) so perhaps he would maintain that the woman can't avow her intention. But if that is how the notion of an avowal is to be understood, it looks as if what manifests a distinctive knowledge of (and control over) intention is not our ability to *avow* these intentions but rather our ability to *affirm* them.

Until now I have been assuming that Moran is using "akrasia" in its standard sense but this may be a misreading. In several places, he is at pains to draw a distinction

⁴ On p. 131 Moran connects claims about alienation to claims about control and responsibility, saying that (certain forms of) the latter are compromised by such alienation (see also pp. 117-8).

⁵ Moran says similar things about the desires which motivate action (pp. 116-20).

between what might be called pure evaluation on the one hand and decision-making on the other (where the notion of “deliberation” applies only to the latter)

the mere appraisal of one's attitudes, however normative, would apply equally well to past as well as to current attitudes, and indeed may have just the same application to another person as to oneself. In itself such an assessment is not an essentially first-personal affair. Rather "deliberative" reflection as intended here is of the same family of thought as practical reflection, which does not conclude with a normative judgement *about* what would be best to do but with the formation of an actual intention *to do* something. (p. 59)

The implication of this paragraph is that Moran is not really interested in the relationship between someone's judgement of their reasons and their action; such judgement interests him only in so far as it involves the formation of an *intention* so to act. What really concerns Moran is the relationship between intention and action. If that is right we should focus our discussion of practical irrationality not on akrasia but rather on what I shall call *irresolution*.

Philosophers sometimes apply the labels “weakness of will” and “incontinence” to both akrasia and irresolution but they are not the same thing.⁷ Akrasia is a matter of failing to intend and act in accordance with your practical judgement, your judgements about what you should (i.e. have most reason) to do. Irresolution is a matter of not sticking to intentions once formed, of giving in to the very ‘inclinations’ which the intention was formed (however akratically) to resist. As Jackson develops his example, his Catholic woman is by turns akratic and irresolute. She is akratic when she forms the intention to attend the abortion clinic but later on her scruples get the better of her and she irresolutely refuses to leave the house on the day of her appointment. We need not take a stand on the issue of whether this woman's akrasia or her irresolution are symptomatic of irrationality. The present point is simply that they are not the same thing.

⁶ This example occurs on p. 4. of F. Jackson – ‘Weakness of Will’ *Mind* (1984) 93, pp. 1-18

⁷ R. Holton – ‘Intention and Weakness of Will’ *Journal of Philosophy* (1999) 96: pp. 241-62.

I suspect that Moran tends to equate akrasia with irresolution (p. 81) and that this goes some way to explaining his view of akrasia. Moran argues convincingly that a person who knows they are irresolute will, as a matter of conceptual necessity, find it very difficult to form intentions (pp. 77-83, pp. 94-8). If past experience of her own behaviour convinces Jackson's woman that she won't attend the abortion clinic even if she decides to do so, it is hard to see how she can even decide to attend the clinic. Saying "I intend to have an abortion but predict that I won't" is rather like saying "I believe that it will rain but it won't". Both sentences express a state of mind that is more than irrational, it is paradoxical.⁸

Of course, an analyst might convince someone that they had a belief which they wished to disavow and the subject might express this by saying "I believe that my brother drove my parents to an early grave even though he didn't". But here the subject seems entitled to enter a qualification: this is not *his* belief in the full sense, he knows about it only via the analyst and he cannot assume full responsibility for it. It is plausible to say the same of intentions that the agent believes they won't execute. The analyst may convince me that I still intend to proposition my childhood sweetheart someday even if I know perfectly well that I'll never bring myself to do it. But, I can plausibly insist, this intention is *mine* only in an qualified sense. Fully self-conscious irresolution is indeed paradoxical. None of this applies to everyday cases of akrasia. There is nothing paradoxical about the statement "I know I ought not to have an abortion but I shall". And there is little pressure to say that such a statement evinces a division of the person, or a diminishment of the person's control over or responsibility for the intention.

Reading Moran's discussion of practical irrationality as concerned with irresolution, he maintains that a person is an authority about what they are doing because they are an authority about the intention with which they act and they are an authority about the intention with which they act because this is something they have created, something which is itself an expression of their agency. Here it is the subject's ability to *affirm* their intention, rather than their ability to *endorse* it, which ensures that they have knowledge of (and control over) what they are doing. But, read in this way,

⁸ This reading makes good sense of the parallels Moran draws between akrasia and self-deception (p.

Moran is not offering the substantial epistemology of agency I suggested he might. Rather he is assuming that an agent has direct knowledge of both of what he is doing and of what he intends to do simply because he is choosing to do it.

To sum up, Moran thinks that there is close tie between what he calls the standpoint of practical deliberation and the possession of first-person epistemic authority: one is an authority about what one is doing and why just in so far as one is occupying this standpoint (p. 127). To occupy this standpoint is, by definition, to both make judgements about what you have reason to do *and* to implement those judgements in decision and action (pp. 63-4, pp. 94-5, p. 131, pp. 145-6).⁹ Now it is clear that these two things can come apart: one can make practical judgements which one does not implement and one can take decisions or perform actions which do not reflect one's practical judgements. The question for Moran is this: when that happens, is first-person authority necessarily compromised? Is a subject who fails to do what he thinks he ought to do *ipso facto* less well placed to know what he is doing and why? My discussion suggests a negative answer.

Conclusion

I have focused on what I take to be a central theme of Moran's book but my treatment has been far from comprehensive. For example, I have had to ignore his interesting discussion of belief, desire and emotion and I have failed to mention his illuminating criticisms of superficially similar positions (like Shoemaker's). Two other features of the book are particularly worthy of mention. First, Moran offers us an interpretation of the work the early Sartre which enabled the present writer to read parts of *Being and Nothingness* with great profit. Secondly, I strongly recommend the fascinating final chapter in which Moran explores some virgin territory: the moral psychology of the first person.

67, p. 87)

⁹ This may be too strong. Perhaps Moran intends only that the deliberative standpoint be one in which it is *believed* that one's decision will reflect one's practical judgement. I don't think this will help. Someone who does not believe but merely hopes that they will take what they regard as the right decision in a difficult situation may still know, in the usual first-person way, what they end up doing (or deciding to do).

Reading a work of academic philosophy, one often finds oneself slogging through chapter after chapter of critical commentary on the work of others, to be rewarded with only the vaguest sketch of an alternative view. Yet (to paraphrase Feyerabend) no sensible person abandons a theory because of a counterexample, only for a better theory. Moran has taken Feyerabend's maxim to heart. His prose is elegant and engaging. He lays out his view in enough detail to expose its weaknesses as well as its strengths. Occasionally, I would have wished to know more about why he thinks his position should be preferred to alternatives. Nevertheless Moran has said enough to encourage others to build up the theory's defences and highlight its advantages over more familiar approaches to the problem of self-knowledge.