

Knowing One's Own Desires

Jonathan Webber

Philosophy of Mind and Phenomenology: Conceptual and Empirical Approaches

edited by Daniel Dahlstrom, Andreas Elpidorou, and Walter Hopp. New York: Routledge, 2016

(To cite by page number, please use the published edition.)

Abstract

Do you know your own desires in some way that other people cannot know them? Richard Moran claims that his influential theory of first-person authority over beliefs and intentions can also cover desires. However, his deliberative model can apply to desire only if one already has some other way of knowing one's own desires. Jean-Paul Sartre's conception of pure reflection, on the other hand, portrays a direct epistemic access to one's own desires that can ground fundamental first-person authority over desires and intentions.

How do you know what you want? Discussions of self-knowledge in contemporary anglophone philosophy tend to focus primarily on how one knows one's own beliefs. This is one aspect of a more general emphasis on the epistemic aspect of the mind. Our status as knowers is subject to far more philosophical scrutiny in this tradition than our status as doers. There is significant attention paid to our knowledge of our own intentions, primarily due to the influence of Elizabeth Anscombe's richly insightful book *Intention*. But this approach to practical agency shares with the emphasis on beliefs a focus on the reasoning aspect of the mind. There is very little discussion of knowledge of one's own desires, even though desire would seem to be closer than intention to the core of what it is to be a creature that seeks to bring about change in the world. This was, by contrast, the focus of a strong tradition of European philosophy whose primary exponents published their greatest works in this area in the first half of the twentieth century.

Richard Moran has drawn on this work of one of these philosophers, Jean-Paul Sartre, in the development of his sophisticated and influential theory of self-knowledge. This theory holds that our epistemic authority to report our own mental states rests on our agential authority to form those mental states. Despite this emphasis on agency in describing the process underlying first-person authority, however, Moran's theory shares with the anglophone tradition that he primarily engages with a strong emphasis on beliefs and intentions as the objects of self-knowledge. Although he intends his theory to apply equally to desire, he provides very little discussion focused on this application of it. This is unfortunate, as we will see, for if we begin the inquiry focusing on the desires at the core of agency as contents of self-knowledge, then there emerges a rather different picture of first-person authority. This alternative account, moreover, is one that is closer to the overall position that Sartre develops in the very same works that Moran discusses.

To see this, we will begin by clarifying Moran's account of first-person authority over belief as consisting in the deliberative authority to form one's own beliefs. We will then consider two ways in which this might be thought to apply to the case of desire, finding that one cannot form desires by deliberating only over whether something is objectively good, but that one can do so by considering whether it is subjectively attractive. This deliberation,

however, rests on already having epistemic access to one's existing desires. We will then turn to Sartre's theory that desires feature in experience by influencing the way the world appears to one. The ability to shift attention from the world itself to the way it seems, which is Sartre's form of phenomenological reflection, thus grounds direct epistemic access to one's desires. We will then see why this form of self-knowledge is the fundamental form of first-person authority over both desires and intentions. Finally, we will see that Moran has overlooked this aspect of self-knowledge in part through focusing on one strand of Sartre's thought at the expense of another that is in tension with it.

1. Moran on the Transparency of Belief

Moran argues that there are two distinct routes to knowledge of one's own beliefs. One route is theoretical reflection, which is inference on the basis of observation of one's speech and other behaviour. This form of self-knowledge is not inherently first-personal, since the very same evidence is available to other observers. Indeed, some of this evidence, such as facial expressions, might be better available to others. The other route is what Moran sometimes calls 'practical reflection' or 'deliberative reflection' (e.g. 2001: 59). If asked whether I believe that p , then in many cases I can answer by directing my attention to whether p is true. This 'transparency' of belief, according to Moran, grounds the inherently first-personal route to self-knowledge. For one can answer questions about one's own beliefs in this way, but not questions about other people's beliefs. What is more, this route to self-knowledge is integral to rational agency itself, since it rests on one's ability to form one's own beliefs. Where this form of self-knowledge fails, one can discern one's beliefs only through theoretical reflection. If a belief is insensitive to one's rational judgment in this way, then one is alienated from it. It is not fully one's own if it is independent of one's endorsement through such a failure of rationality.

In his book *Authority and Estrangement*, where the theory is developed most fully, Moran often explicates this idea of transparency as the question of whether one believes that p being answerable by addressing the question of whether one is to believe that p , or has good reason to believe that p . Thus, for example, 'a person is credited with first person authority when we take the question of what he *does* believe to be settled by his decision as to what he *is to* believe' (2001: 134). Moran's idea here is that if one deliberates about what one is to believe, one must be assuming that this will decide one's belief (2001: 94-5). For

this reason, Moran sometimes describes the question of whether one believes that p as transparent, from one's own perspective, to the question of whether p is true (2001: 74-7, 85). This raises the question of how we should understand the relation between considering whether p is true and considering whether one should believe that p . Which should one directly consider in order to exercise first-person authority over one's beliefs?

There are good reasons to reject the idea that the question of what one believes is directly transparent to the question of what one should believe. For one might judge that one has good reason to believe that p without thereby believing that p . For example, one's reasons might be 'good enough to permit belief in p , without requiring belief in p , and thereby also permitting suspension of belief in p ' (Way 2007: 228). If one exercises first-person authority by considering what one has good reason to believe, then in this case one would have the authority to claim to believe that p even if one had suspended belief whether p . Moran could not classify this as a failure of first-person authority due to a failure of the rationality that grounds that authority, since this suspension of belief is rationally permitted. Moreover, the question whether one has good reason to believe that p itself has a reflective aspect. If the question of what one believes is transparent to the question of what one has reason to believe, then we need an account of how one finds out what one's reasons for belief are. This cannot be a matter of observational self-knowledge, for if it were then the relevant information would be publicly available and so would not ground a form of first-person authority. What is required is a further form of self-knowledge. If there is such a form of self-knowledge, it might undermine the motivation for Moran's account of non-observational self-knowledge by fulfilling the role for which he proposes that account (O'Brien 2003, 379).

Moran's theory is therefore best interpreted as holding that the question of whether one believes that p is directly transparent to the question of whether p is true, with his emphasis on having reasons to believe that p indicating the practical procedure that one needs to adopt for rationally deciding whether p is true (O'Brien 2003: 379). To put the point another way, the question about what one believes is only indirectly transparent to the question about what one has reason to believe, where one should consider one's own reasons only in the service of the prior question of what the truth is. Moran seems to endorse this reading that accords the truth of p priority over reasons for believing p when he writes, in a paper discussing his book, that transparency means that 'a person answers

the question whether he *believes* that p in the same way he would address himself to the question whether p itself' (2004: 457; see also 2004: 468; 2012: 235).

It follows from this that we should not accept Moran's claim that the transparency 'of first-person belief reports' has its 'source in the primacy of a deliberative rather than a theoretical stance towards one's own state of mind' (2001: 64). It rather has its source in deliberation about the world, which does not require any sort of stance towards one's own state of mind. It is likewise misleading to describe this process as 'deliberative reflection', since it does not essentially include oneself or one's mind in its content. It is true that one can announce the result of the deliberation either by ' p is true' or by 'I believe that p ', because the judgment that p makes it the case, at that time at least, that the judge believes that p . It is also true that this avowal of one's own belief is a reflective formulation of the affirmation of its embedded proposition. But this reflectivity belongs to the report, not to the deliberative process that resulted in it. Only one of the two routes to knowing one's own beliefs that Moran discusses, therefore, is a form of reflection. This is the 'theoretical reflection' that relies on evidence that is available to other observers as well as to oneself.

2. Desire and Desirability

What leads Moran to slip between the question of the truth of p and the question of whether to believe that p ? Perhaps it results from his aim of formulating a unified account of transparency across mental states of different kinds. If asked whether I intend that p , for example whether I intend that I go to the cinema this evening, then I cannot answer this by considering whether p is true, for to do that would be to make a prediction rather than form an intention. Thus, it is more plausible to say that the question whether I intend that p is transparent to the question whether to intend that p , rather than to the question whether p is true (Way 2007: 225). Questions about my beliefs and my intentions differ, therefore, in the questions to which they are transparent for me. But their properties of transparency are nevertheless unified at a higher level of generality. For in both cases, if these claims are correct, one can answer the question about one's mental state through the deliberation that forms that mental state.

What about desire? If questions about what one wants are also transparent, then what is the nature of the deliberation whose conclusion can at once form and announce one's

desire? It may be objected that desires are not generally formed in this way. But the theory does not require that they are. Neither does it require that beliefs or intentions are generally formed by deliberation. It requires only that one can answer a question about these mental states through deliberation whose conclusion forms the mental state that it announces. It might still be objected, however, that we have first-person authority to report desires that we feel but which cannot be formed or reshaped by deliberation. Moran's strategy for dealing with this objection is to distinguish between mere bodily desires and judgment-sensitive rational desires (2001: 114-8). He does not provide an account of how we know our bodily desires, but since he sees these as impinging on the rational mind from without we can assume that he considers our awareness of them to be a form of sensation or perception (2001: 114-5).

There are good reasons not to accept Moran's distinction between two kinds of desire (Webber 2016: §§ 3-4). For present purposes, however, we can formulate our question solely in terms of those desires that Moran considers to be sensitive to deliberative judgment. What is the subject primarily aiming to discern through the deliberation to which these desires are purportedly sensitive? It cannot be the truth of the content of the desire. My desire that p cannot be formed by deliberation about whether p is true. It must rather be formed through deliberation about whether p is desirable, as Moran acknowledges (2001: 57, 60). So if I am asked whether I want to go to the cinema this evening, if this theory is correct, I cannot answer by considering whether I do go to the cinema this evening, but I can answer by considering whether my going to the cinema this evening is desirable. The judgment produced by this deliberation would both form and announce my desire.

How should we understand the notion of desirability in play here? It could be read objectively, to mean that the deliberation is concerned with whether p is or would be good. This would parallel the idea that one's belief about p can be formed and announced through deliberation about whether p is true. In both cases, the deliberation would make no essential reference to the subject's own mental states. It would rather be wholly concerned with some aspect of the world beyond that mind. For this reason, deliberation about whether some possible state of affairs is or would be good is not a form of reflection. One's deliberative conclusion could be announced in the form of the avowal 'I want that p ', but this reflectivity would be a feature of the avowal rather than the deliberation.

Moreover, this might be thought to embody a normative feature of desire that parallels a normative feature of belief. It might be argued that one ought to desire what is good just as one ought to believe what is true, so these forms of transparency reflect a normative demand of rationality.

Alternatively, however, we could understand the deliberation to concern desirability in the subjective sense. That is, it might be deliberation about whether p is attractive rather than about whether p is good. I would determine whether I desire something by deciding whether I have good reason to desire it. The transparency of the question of one's own desire would parallel that of the question of one's own intention. This practical deliberation, moreover, would be a form of reflection, since it would need to take into account one's existing motivational states. In order to decide whether to intend that p , one would need to consider how p would fit with one's other desires and intentions. Likewise, to consider whether p is attractive to one requires considering its relation to one's existing desires. Should we understand the purported transparency of desire to parallel that of intention in this way? Or should we understand it in the way that parallels the transparency of belief?

3. The Desirable As The Good

Moran's theory of first-person authority requires that deliberation culminate in a judgment that forms the relevant mental state. It does not require that deliberation always or necessarily has this effect. But it does require that this should be the normal result of deliberative judgment, with other cases being failures of rational functioning. For otherwise one would not in general be warranted in answering a question about one's mental state by forming a judgment about the object of that mental state. Deliberation about p could warrant the avowal 'I desire that p ', for example, only if judgments arrived at through that kind of deliberation generally had the effect of forming one's desires. For this reason, it seems that first-person authority over whether I desire that p cannot be grounded in deliberation over whether p is good. The judgment that p is good does not in general result in the desire that p , the judgment that p is not good does not in general result in the absence of desire that p , and the judgment that p is bad does not in general result in the desire that not- p .

The most obvious examples of this rational insensitivity of desire to judgments of goodness are ones that are bound up with bodily needs, whether these are like hunger and thirst in being intrinsic to the functioning of the human body or whether they are like the smoker's desire for a cigarette in being an acquired appetite. But this insensitivity is a feature of desires more generally. One might decide that it would be good to start cycling to work instead of driving, but still want to drive. One might judge that spending the day rewatching old movies would not be a good use of one's time, but want to do it anyway. One might decide that speaking one's mind to one's boss on this occasion would be a bad move, but still want to do so. Philosophers have puzzled over exactly how one can end up doing something other than what one judges to be good. But the point here does not concern akratic action. The point is just that it is quite common to continue to have a desire that runs contrary to one's judgment, whether or not one then acts on that desire.

Can these cases be understood in terms of the distinction between desiring the object and desiring an aspect of the object? It certainly can happen, for example, that I judge that cycling is preferable to driving when all things are considered, but nevertheless still want the easier physical activity of driving. But this distinction does not capture every way in which desires that are not rooted in bodily needs might persist despite contrary judgments. For it may not simply be a desire for an aspect of driving that survives my judgment that cycling is better. It may instead be the desire to drive. That is, the desire for the object, with all of its aspects, survives the contrary judgment about that object in light of all of its aspects. This may perhaps be explained by the relative weights I place on the various aspects of the object in deliberation differing from the relative strengths of my desires for those aspects. But whatever the explanation, it seems that the object of the judgment does match the object of the desire.

In discussing his book, Moran provides an example where the judgment and the contrary desire concern the same aspect of the object, rather than merely the same object. His example is judging some activity to be unworthy of desire due to its disgusting character. Not only is it true that a desire for the activity can survive this judgment, but moreover 'surely some activity's disgusting character can be just what appeals to me and makes the pursuit of it alluring, even rewarding, even if it conflicts with other values of mine' (2004: 472). This is not a case where one judges the activity good despite finding it disgusting, but rather a case of wanting something for those aspects of it that one judges to be bad. Moran

does not classify this as a case of rational malfunction. If he did, then he would need to claim that first-person authority over desire fails in this case, that one can only learn of the desire through observation, and that one would thus be alienated from the desire. Rather, he agrees that these implications are not right. It is, he writes, ‘a desire for that thing *because* it is both trivial and disgusting, something I desire under that very aspect; and yet I still acknowledge it as my own’ (2004: 472).

How is it possible to have first-person authority in such a case, on Moran’s account? He does not give a clear answer to this. He does agree that ‘what is said about belief will only carry over’ to this kind of case with ‘substantial modifications’ and that ‘any connection between desire and being found worth desiring has to be interpreted very broadly to be psychologically realistic at all’ (2004: 471). But he does not say any more about what these modifications or this breadth would consist in. What these examples show is that we should not consider normal psychological functioning to bring desires immediately into line with judgments of goodness. Perhaps creatures with a different form of rationality, perhaps even creatures with a more perfect form of rationality, would be responsive to the good in this way, but our form of rationality is not like this. So if the question of whether I desire that p is answerable through deliberation about the desirability of p , then that desirability must be attractiveness to me rather than objective goodness.

4. The Desirable As The Attractive

Moran’s example of a desire being sensitive to rational deliberation is one in which the deliberative judgment concerns the attractiveness of the object. ‘When someone wants to change jobs’, he argues, this desire ‘depends for its justification on various beliefs (about oneself, about one’s present job and prospects elsewhere)’ and it is a normal expectation that changes in these beliefs that render the desire unjustified will result in no longer having the desire (2001: 115). What are the beliefs on which the desire rationally and psychologically ought to depend? Moran says only that judgment-sensitive desires depend on ‘beliefs about what makes these various things desirable’ (2001: 115). That is, not only factual beliefs about the object of desire, but also beliefs about the desirability of the aspects of that object picked out by these factual beliefs. We have already seen that this desirability cannot be understood as objective goodness, since it is implausible to suggest that desire is normally dependent on judgments of goodness. We should rather see the

judgment as dependent on the beliefs about how attractive or unattractive each aspect under consideration is.

Since the attractiveness of the object as a whole is a matter of the attractiveness of each of its aspects, deliberation over whether something is desirable in this sense concerns not only the object itself but also one's own existing mental states. It is properly described as a reflective process, because the reasoning concerns oneself, or at least one's own desires, as well as the object under consideration. The desire to change jobs is dependent not only on beliefs about the nature of one's current job and the nature of other jobs one might have instead, but also about how much one wants or does not want a job with these or those aspects. Thus, one can answer a question about whether one desires that p by considering the various aspects of p in relation to one's desires and thereby forming a judgment about the overall attractiveness of p itself.

If this is right, a question about one's own desire with regard to some object is transparent to a question about the overall attractiveness of that object, which one is to answer through deliberation that considers the various aspects of that object in relation to one's desires. This would not reflect a strong normative demand, if indeed there is one, that one's desires ought to track the good just as one's beliefs ought to track the truth. But it would nevertheless reflect a weaker normative demand that one ought not to desire some object overall if achieving that object would frustrate one's desires more than it would satisfy them. Such a demand is not so strong that it requires consistency of desires. Yet it is strong enough to ground the expectation that one's desire be sensitive to changes in information one has about the object of that desire.

Obviously, this deliberation about the object requires one to have knowledge of one's own desires. One cannot consider whether an aspect of the object meets one's desires unless one knows what the relevant desires are. If this is how we are to understand deliberative knowledge of one's own desires, then this is not a fundamental form of self-knowledge, for it rests on prior knowledge of one's own desires. Moran's account of self-knowledge only allows for one possible source of this prior knowledge, which is observation. But there are two good reasons to reject the idea that deliberative first-person authority over one's desires rests on observational knowledge of one's desires. One reason is that this would seem to undermine the claim that the transparency of desire to attractiveness grounds a

distinctively first-personal form of self-knowledge. If the deliberation simply involves making calculations on the basis of information about my desires that is in principle available to any observer, then in principle anyone can deliberate in this way about whether that object is attractive to me.

The second reason is that the ambiguity of the observational evidence would undermine any normative demand not to desire outcomes that would have a negative net effect on the overall satisfaction of one's desires. This is well illustrated by a vignette in Sartre's novel *The Age of Reason* in which Daniel wants to prove that he is not the sentimental person that other people take him to be. His relationship with his cats is central to the evidence on which this view of him is based. So he forms the desire to drown his cats in the river Seine. When he gets to the water's edge, he finds that he cannot bring himself to do it. He had not taken account of his strong desire to continue caring for his cats. Yet he had taken account of the observational evidence, which he interpreted in a way that did not indicate this strong desire (AR: 81-91). Observational evidence is generally, perhaps necessarily, open to more than one interpretation. So how could there be a normative demand to take one's desires into account if one could not reliably identify them?

5. The Phenomenology of Desire

If desires are to be taken into account in rational decision making at all, then they must feature in experience in some way. For they could not motivate rational behaviour unless the subject were sufficiently aware of them to take them into account in the reasoning behind that behaviour. This is not to say that desires do not motivate behaviour in other ways. In some cases, for example, they might subtly direct reasoning to the desired conclusion through influencing the rational process itself rather than through being or shaping the considerations taken into account. But this would seem to be paradigmatic of irrational decision making. The normal role of desire in rational agency must be to contribute to deliberation in a way that does not undermine the rationality of that process. In order to play this role, they must feature in the deliberator's experience in some way.

Moran does not consider in any detail just how desires feature in experience and thus feed into rational deliberation. He does describe 'brute' desires, such as 'those associated with hunger or sheer fatigue' as 'experienced by the person as feelings that simply come over

him' and as things that 'must simply be responded to' much like 'any other empirical phenomena' (2001: 114-5). This idea that some desires are experienced as sensations might well be correct of extreme cases of these bodily desires, though as we will see the same desires in less extreme form are experienced in another way. But this cannot account for the way desires generally feature in experience. Moran rightly distances his view from the idea that all desires, including for example the desire to change jobs, 'simply assail us with their force' (2001: 116). His only other model of our relationship with our desires, however, is the deliberative model that we have seen to require some prior awareness of desire. So we are left with the question of how desires generally feature in experience.

Sartre took this question to be central to understanding human motivation. We do not generally experience our desires as inner drives or forces, he argues, but rather as the directive structure of the world we experience. The desire to get to work on time, for example, is experienced in the sound of the alarm clock being experienced as a demand to get out of bed rather than as a mere annoyance (B&N: 62-3). Daniel is confronted with his own desire to continue caring for his cats when he finds he cannot bring himself to throw them into the river. The desire not to deal with slimy things is manifested not as sensation of disgust in their presence, but as those things themselves seeming to have the property of being disgusting (B&N: 630-2). Similarly, thirst can be experienced in the glass of water looking inviting, hunger in the plate of food seeming attractive, and tiredness in the difficulty of simple tasks. This account is intuitively plausible. We do not experience the world as an arena of physical properties to which our desires are subjective reactions. We experience it as already being appealing, exciting, forbidding, repulsive, and so on. But objects in the world can only have these properties in relation to our existing motivations.

On this view, unreflective rational guidance of action is a matter of responding to this field of reasons for action, which in turn has been shaped by our desires (B&N: 472-3). This grounds an account of how we can become reflectively aware of our desires. If we wish to turn attention to the question of what we desire, we need only shift our focus from the world of objects we experience to the experience of that world. Since the experience is nothing but the presentation of the world as having a particular directive structure, this shift of attention does not provide access to some inner mental states that represent the world. Rather, it brings into focus the directive structure that the world is experienced as having. Daniel can turn his attention away from his cats and focus instead on the appeal

that they have for him. I can shift my attention from the alarm clock sound itself to the insistent demand it appears to make. This shift brings one's desires to light.

Cats and alarms are concrete perceivable objects, but the same point can be made about our desires in relation to more abstract things. In order to decide whether I want to change jobs, I must not only identify the various aspects of my current job and those of other jobs I might realistically have, but also identify how these aspects relate to my desires. I can do so by thinking about those aspects and considering how they strike me. When I think about my current commute, does it seem arduous? Does the prospect of spending the day with my colleagues seem inviting? The same can be said of Moran's other example. I know that it is the trivial and disgusting nature of the activity that attracts me because I find that particular triviality and disgustingness appealing when I think about it. Since our desires shape the way their objects appear to us, then, we can access those desires by considering those appearances.

6. Phenomenological Reflection and First-Person Authority

Directing attention on one's experience in this way is not well described as 'introspection'. It does not involve looking inside anywhere. Rather, one's attention continues to be directed towards the world. But this attention is no longer focused on the object itself, but rather on the way it seems to be. When I run after the tram, for example, my attention is focused on the tram itself, but I can also shift attention to the appearance of the tram as something-to-be-caught. In doing so, I do not direct my attention inwards: it remains directed outwards, towards the tram, but shifts its focus from the tram itself to how the tram seems. This is what Sartre describes as 'pure reflection' and since it is a form of bracketing the reality of the object to focus on its appearance, bracketing the being to focus on the phenomenon, it is also rightly described as a form of the phenomenological reduction. But this is not a technical philosophical procedure and neither is there anything mysterious about it. It is rather an ordinary part of everyday life.

This self-awareness is a form of transparency. The idea here is not that experience itself is transparent to the world, presenting nothing other than its objects. Indeed, the transparency involved in self-awareness rests on that not being true. Sartre describes consciousness as 'translucent' (*translucide*), not transparent (B&N: 67, 72, 78, 101, 119, 357,

409). The difference between these terms in French is the same as in English: whereas one can see objects plainly through a transparent object, like an ordinary window, a translucent object distorts the way those objects seem. Our motives distort the appearance of things. It is this translucency of our experience of the world that allows for the logical transparency, in Moran's sense of the word, of questions about one's own desires to questions about how the world appears to one. To answer the question about whether one desires that p , one needs to direct attention to the prospect of p , either as a whole or in its various aspects, and report on how that prospect seems to one.

Moreover, this phenomenological reflection meets the criteria that Moran lays down for an acceptable theory of first-person authority (2001: 32-5). First, it is inherently first-personal. The way the world seems to you is something that is directly epistemically accessible only to you. I might infer it from your behaviour, as indeed can you, but this is an indirect access to the information. Or I could ask you how things seem to you, but this rests on your already having access to that information. Only you can be directly aware of how things seem to you. Second, it is clear why we only have this access to some aspects of ourselves and not others. Since phenomenological reflection is attention to the way the world seems in experience, it can provide access only to states of ourselves that are manifested in that seeming. This can include some physical states such as hunger, thirst, and tiredness, as well as mental states such as desires, but would generally exclude the details of our health. Moreover, in focusing on desire rather than belief and intention as its central case, this is an account of epistemic access to what is arguably the core of mindedness, the core of being an agent who seeks to change things.

Most importantly, this account shows first-person authority to be integral to rational agency rather than something additional to it. Moran's account of first-person authority over belief rests on the ability to make up one's own mind through deliberation. Moran points out that this puts rational agency at the centre of the account (2001: 150-1). But this is only because this first-person authority supervenes on rational agency. That is, the ability to express the deliberative judgment as an avowal of belief does not itself seem essential to the ability to form beliefs through deliberative judgment or to the ability to express beliefs in words and actions. The phenomenological account of first-person authority over desire, by contrast, has the implication that this epistemic authority is essential to rational agency. For one cannot rationally shape one's desires without regard to one's other desires.

Deliberation about the desirability of some outcome rationally ought to take into account one's existing desires.

Should we consider this self-knowledge to be another kind of first-person authority alongside the deliberative kind that Moran describes? Once an account of first-person authority is available that covers cases that Moran's cannot cover, the question arises whether it can be extended to displace Moran's (O'Brien 2003: 379). Deliberative first-person authority over desire rests on this phenomenological reflection on one's existing desires, so inherits its status as authoritative from it. A similar point can be made about knowledge of one's own intentions. When he formed the intention to drown his cats, Sartre's character Daniel should have reflected on his desires concerning those cats. One's intentions are not authoritative unless formed with due regard to one's existing desires. But we have already seen that the question of what one believes differs in being directly transparent to the truth, rather than transparent to what one has reason to believe. It might turn out, on further consideration, that deliberation about the truth rests on something parallel to phenomenological reflection on desire. For now, however, we should conclude only that rational agency in the practical sphere rests on phenomenological reflection.

7. Two Strands of Sartre's Thought

Moran develops his account of first-person authority in part through his analysis of Sartre's example of the akratic gambler (2001: 78-83; B&N: 56-7). This gambler has resolved not to gamble again, but that does not help him to resist the charms of the casino in front of him. Moran argues that this is because the theoretical perspective that I can take towards my own intention, or resolution, as grounding a prediction of my behaviour depends on my own deliberative relation to the content of the same intention or resolution. 'I must recognise', writes Moran, 'that the resolution only exists as a fact on which anyone can base a prediction insofar as I continue to endorse it' (2001: 82). When the gambler feels his commitment to not gambling is not strong enough, the reason that turning to his resolution for support will not help him is that the resolution itself is only as strong as his current commitment to not gambling.

This is a subtle and illuminating account of Sartre's example. Moran uses it to develop the idea that intentions are sensitive to deliberative judgment in the way that grounds his

theory of first-person authority. He extends the case by drawing on Sartre's thoughts about ascribing beliefs to oneself (B&N: 92-3). Why would I treat my belief itself, such as my belief that Pierre is my friend, as evidence for the truth of that belief? I would do so only if I am no longer confident about whether the content of that belief is true and I cannot find sufficient reasons for it elsewhere. But this motivating state itself means that the belief cannot provide good evidence, since the fact that someone – whether myself or anyone else – unconfidently believes that Pierre is my friend does not provide me with good reason to believe that Pierre is my friend (2001: 83). However, this parallel reasoning concerning intentions and beliefs does not consider whether the same can be said about desire or what difference a consideration of desire would make to the case of intention.

Sartre's discussion of his example similarly overlooks desire. Why is the gambler tempted when faced with a casino? Because he has a persisting desire to gamble. This grounds an important disanalogy between intentions and beliefs. When forming intentions, one ought to take into account not only one's existing intentions but also one's desires. Deliberation resulting in belief, on the other hand, needs to be sensitive only to one's other beliefs, not to one's mental states of any other kind. For this reason, it can seem odd that Sartre describes the akratic gambler as discovering his radical freedom through his awareness that his prior intention cannot help him to resist temptation. The gambler has resolved to resist his desire and then finds that the desire persists while the strength of his resolution does not, just as Daniel is confronted with his recalcitrant desire to care for his cats. This is why Sartre's example can look rather more like an illustration of the limitations of freedom. It looks as though the lesson to be learned is that the gambler should have taken his desires into account more carefully when making his decision, with the implication that he is no longer free to choose to give up gambling.

This oddness is due to the structure of *Being and Nothingness*, which is not a series of arguments for discrete conclusions but rather the progressive elaboration of an overall philosophy none of whose aspects are fully articulated until the elaboration is complete. Sartre goes on to argue that one's desires are themselves dependent on one's deliberative judgment. Desires are grounded in the projects that I pursue, or more precisely in the values at their core, he argues, and persist only while I endorse those values and so remain committed to those projects (B&N: 459; Webber 2009: chapters 3-5). Sartre sometimes seems to hold that we can shape our desires through deliberative judgment about the

desirability of their objects, even though this would bring with it an alteration in the projects in which they are grounded (B&N: 486, 497-8). But in other passages, he seems to hold that we have only indirect control over our desires through deliberation over our projects, values, or their objects, which is why we can need the help of an existential psychoanalyst to uncover the project in which a desire is grounded (B&N: 495, 591). Either way, the gambler is not constrained by his desire to gamble, even though he has failed to consider just how strong that desire is and thus just how difficult it will be for him to abandon the old project of gambling.

Ultimately, this vacillation is due to the tension between two strands of thought in *Being and Nothingness*. One strand elaborates the idea that my mental states have no inertia of their own, but persist only with my endorsement. The other develops the idea that the reasons for action that I find in the world reflect the projects I have already undertaken. Moran focuses almost exclusively on the first of these (see e.g. 2001: 140). This distorts his picture of the formation of intention by ignoring the apparent inertia of desire, a distortion which occludes the differences between the theoretical attitude of belief and the practical attitudes of desire and intention. This is what has led Moran to an account of first-person authority in general that cannot be fundamental in the case of the practical attitudes. The idea that first-person authority over desires and intentions is rooted in phenomenological reflection, on the other hand, is consistent with each strand of Sartre's thought. But it does not require Sartre's particular elaborations of either strand and Sartre's vacillation concerning deliberative control of desire suggests that the two strands might not be mutually consistent. Whether that suggestion is correct and what difference it makes to first-person authority, however, are questions for another time.¹

¹ This paper was developed through presentations at Cardiff University work-in-progress seminar, Kings College London philosophy society, the visiting speaker seminar at Manchester Metropolitan University, and the 2013 conference of the Nordic Society for Phenomenology at University of Copenhagen. I am grateful to the organisers and participants of these for helping to shape my thoughts on this issue. I am also grateful to the editors of this volume for helpful feedback on the first draft of the paper.

Works Cited

- Anscombe, G. E. M. 1957. *Intention*. Oxford: Basil Blackwell.
- Moran, Richard. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- Moran, Richard. 2004. Replies to Heal, Reginster, Wilson, and Lear. *Philosophy and Phenomenological Research* 69: 455-472.
- Moran, Richard. 2012. Self-Knowledge, 'Transparency', and the Forms of Activity. In *Introspection and Consciousness*, edited by Declan Smithies and Daniel. Oxford University Press, 2012.
- O'Brien, Lucy. 2003. Moran on Agency and Self-Knowledge. *European Journal of Philosophy* 11: 375-390.
- Sartre, Jean-Paul. AR. *The Age of Reason*. Translated by Eric Sutton. Penguin: Harmondsworth, 1986. First published as *L'Âge de Raison*, Paris: Gallimard, 1945.
- Sartre, Jean-Paul. B&N. *Being and Nothingness: An Essay in Phenomenological Ontology*. Translated by Hazel E. Barnes. Edited by Arlette Elkaïm-Sartre. London: Routledge, 2003. First published in 1943. Translation of *L'Être et le Néant: Essai d'Ontologie Phenomenologique*, revised by Arlette Elkaïm-Sartre, Paris: Gallimard, 1994. Original edition published in 1943
- Way, Jonathan. 2007. Self-Knowledge and the Limits of Transparency. *Analysis* 67: 223-230.
- Webber, Jonathan. 2009. *The Existentialism of Jean-Paul Sartre*. New York: Routledge.
- Webber, Jonathan. 2016. Habituation and First-Person Authority. In *Time and Agency*, ed. Roman Altshuler and Michael Sigrist. New York: Routledge.