

Self-Knowledge, Deliberation, and the Fruit of Satan

Josep E. Corbí¹

Received: 30 October 2015 / Accepted: 20 July 2016 / Published online: 13 August 2016
© Springer Science+Business Media Dordrecht 2016

Abstract Robert Dunn (2006) and Richard Moran (2001) have emphasized the importance of deliberation to account for the privileged authority of self-ascriptions. They oppose a theoretical attitude toward oneself to a deliberative attitude that they regard as more intimate, as purely first-personal. In this paper, I intend to challenge Dunn's and Moran's understanding of how the deliberative attitude is to be conceived of and, in particular, I will call into question their claim that this attitude is wholly non-observational. More positively, I will elaborate on the sort of self-observation that must play a central role in an agent's deliberation if she is to recognize a certain belief, decision, or intention as genuinely her own and, therefore, as expressing a purely first-person point of view. In the elaboration of my argument, I will rely on a number of situations as they are described in Peter Carey's novel *Oscar and Lucinda*.

Keywords Self-knowledge · Expression · Self-observation · Deliberation · Internalization · Awareness

Robert Dunn (2006) and Richard Moran (2001) have emphasized the importance of deliberation to account for the privileged authority of mental self-ascriptions.¹ They oppose a theoretical attitude towards oneself, which is merely observational and provides evidential basis for any such self-ascription, to a deliberative attitude that they regard as more intimate, as purely first-personal, so that self-ascriptions of beliefs, decisions, or intentions in the context of this attitude are particularly authoritative even though they are not grounded on any specific evidence about oneself. In this paper, I intend to challenge Dunn's and Moran's understanding of how the deliberative attitude is to be conceived of and, in particular, I will call into question their claim that this attitude is wholly non-observational and, therefore, that self-observation cannot form a part of a purely first-person perspective. More positively, I will argue that a certain sort

¹A similar model has been sketched by Boyle (2009), Burge (1996), and Zimmerman (2008).

✉ Josep E. Corbí
josepcorbi@gmail.com

¹ University of Valencia, Valencia, Spain

of self-observation must play a central role in an agent's deliberation if she is to recognize a certain belief, decision, or intention as genuinely her own and, therefore, as expressing a purely first-person point of view. In the elaboration of my argument, I will examine a number of passages in Peter Carey's novel *Oscar and Lucinda*. Such passages combine a meticulous description of a situation that a certain character faces and the slight bodily changes, experiences, thoughts, and actions that come as her response to it. All these elements tend to form an intricate pattern that highlights how hard it is to shape a point of view that could be identified as genuinely one's own.

The paper is structured as follows. In section 1, I will introduce Dunn's and Moran's deliberative model of self-knowledge in some detail. In section 2, I will vindicate the normative authority of the deliverances of a certain kind of self-observation and introduce the notion of "expressive self-awareness" to name it. In section 3, I will argue that Dunn's own account of evaluative statements forces him to acknowledge the indispensability of this kind of self-observation. In section 4, I will show how an agent's deliberation is often inadvertently shaped by someone else's views and rely on the notion of expressive self-awareness to elucidate the conditions under which a view that an agent holds can be recognized as genuinely her own. This is, as I see it, the ultimate purpose of Dunn's and Moran's account of self-knowledge, namely: to individuate a kind of attitude towards oneself that could be recognized as purely first-personal insofar as it is closely connected to an agent's capacity to lead a meaningful life. My point is, however, that, unlike what they claim, a certain kind of self-observation must be included within the deliberative attitude if the latter is to apprehend what counts as a genuinely first-person point of view.

1 The Deliberative Model of Self-Knowledge

Both Dunn and Moran seek to account for the privileged epistemic authority of certain self-ascriptions; in other words, they try to explain how it is that the epistemic authority of some self-ascriptions are enhanced, rather than diminished, by the fact that they do not rely on any specific evidence:

"What remains before us, then, is a basic asymmetry between first-person and third-person relations. A person can make reliable psychological ascriptions to himself immediately, without needing to observe what he says and does... But perhaps more pressing than the question about the epistemic completeness or reliability of introspection are philosophical questions concerning how there could even *be* such a thing as this capacity, however imperfect its deliverances." (Moran 2001, p. 12. See Dunn 2006, pp. 37, 38; Moran 2001, p. 32)

For this purpose, they explore two sorts of attitudes that an agent may adopt towards herself. To illustrate these two attitudes, Moran examines the case of an akratic gambler whom, according to Sartre's story (Sartre 2003: part II, sec. 1), *decides* at some point to stop gambling and almost simultaneously *predicts* that he will gamble again:

“Sartre’s case if the akratic gambler who resolves to stop gambling is in some ways a more helpful example for considering the two stances and the contrasting roles of *commitment* (of oneself) and *theoretical knowledge* about oneself. For the gambler to have made such a decision is to be committed to avoiding the gaming tables.... But now, at the same time, he does know himself empirically too; he knows his history, and from this point of view his ‘resolution’ is a psychological fact about him with a certain degree of strength.” (Moran 2001, p. 79; my emphasis)

Moran stresses, however, that our agency cannot intelligibly renounce either attitude.² To elaborate on the distinction between the theoretical and the deliberative attitudes, let us examine how the following three questions interrelate:

- (Q1) Do I desire p?
- (Q2) Is p good?
- (Q3) Am I to desire p?

According to Dunn and Moran, question (Q1) can be interpreted in two ways.³ There is, to begin with, a *theoretical approach* to it, so that the agent answers question (Q1) by *observing* her behavior and experiences the way the akratic gambler gathers evidence about himself in order to predict that he will gamble again. To this end, he examines his own psychological condition from a *detached, uncommitted perspective*, that is, he approaches his own desires as *non-evaluative facts* about his psychological dispositions that he has discovered in light of some amount of *evidence*. From this point of view, question (Q1) is quite alien to questions (Q2) and (Q3), for the agent’s answer to question (Q1) provides no immediate grounds to address those normative or evaluative questions in one or another way. This is not to deny, of course, that the non-evaluative facts thus discovered might have some *instrumental* value relative to an independent decision or commitment that the agent might have and, in this limited sense, they will be relevant to question (Q3). We must, though, oppose the instrumental significance of non-evaluative facts to a sort of normative import or authority that some experiences, features, and situations may have and such that they do not depend on any particular goals that the agent might actually pursue. We could thus ascribe to these experiences, features, and

² In this respect, Moran claims: “There is one kind of evasion in the empty denial of one’s facticity (e.g., one’s history of weakness and fallibility), as if to say ‘Don’t worry about my actual history of letting you down, for I hereby renounce and transcend all that.’ But there is also evasion in submerging oneself in facticity, as if to say, ‘Of course, whether I will in fact disappoint you again is a fully empirical question. You know as much as I do as to what the probabilities are, and so you can plan accordingly.’” (Moran 2001, p. 81)

³ In fact, Dunn and Moran defend a *hybrid* model of self-knowledge (Gertler 2011, p. 167), since they both assume that a theoretical attitude towards oneself also delivers self-knowledge. It is essential to their view, however, that these two varieties of self-knowledge are not on equal terms, since only the deliberative attitude allows us to articulate a purely first-person point of view. This is why Gertler herself refers to this model as the rationalist theory as opposed to an empiricist view (Gertler 2011, Chaps. 6, 8).

situations a *prima facie normative authority* that, of course, some other concerns, experiences, or situations may eventually challenge, cancel, or counteract. The theoretical attitude, as it has been conceived of, cannot intelligibly deliver experiences or situations with this kind of normative import. The main point in this paper is, however, that there is more to self-observation than the theoretical attitude can provide, as we shall see in sections 2 to 4.

On the *deliberative interpretation*, question (Q1) appears as *transparent* to question (Q2), that is, those *reasons* that would support a positive or a negative answer to (Q2) would also motivate the corresponding answer to (Q1) and, derivatively, to (Q3). Thus, we can say that, on this second interpretation, (Q1), (Q2), and (Q3) are *deliberative transforms* of each other and, in this respect, they all meet the *transparency condition*.⁴ Dunn and Moran regard this *deliberative* attitude towards one's desires as constitutive of a purely first-person attitude, whereas the theoretical attitude towards oneself appears as essentially similar to that of a third party. A seemingly trivial implication of this line of reasoning is that self-observation plays no role in a deliberative answer to question (Q1) and this is why the deliberative attitude towards one's own desires is presented as *wholly non-observational*.⁵

It is clear, however, that the deliberative attitude towards one's desires would be of scarce use to lead one's life if it didn't typically have an impact on our desires as they are identified from a theoretical perspective. The assumption is that an agent's psychological dispositions are typically *permeable* to her decisions and commitments and, therefore, to her answer to questions (Q1)–(Q3) inasmuch as they are conceived of as deliberative transforms of each other. Dunn and Moran do not present this permeability requirement as a merely contingent link, but as *constitutive* of our agency. In fact, a lack of permeability emerges as an essential feature of some kinds of psychic impairment, like that of the akratic gambler. Psychoanalytic cases are interpreted similarly. The agent may be able to accept the analyst's interpretation of her neurotic symptoms and,

⁴ The transparency condition was first introduced by Gareth Evans with regard to beliefs and, in particular, to highlight how we do typically answer the question "Do I believe that P?" by exploring a world and, therefore, by looking for an answer to the question "Is P true?" It follows that whatever reason an agent may have to answer the latter question in a certain way will also ground a similar answer to the former one (Evans 1982, p. 225). Dunn and Moran defend the view, however, that a similar procedure applies to desires as well, that is, they argue that there is a way of raising the question "Do I desire P?" that calls for an examination of the question as to whether P is worth-desiring rather than for an exploration of one's actual psychological dispositions.

⁵ As Dunn repeatedly puts it: "Some first-person self-knowledge is wholly observational, as when I am consciously in pain. Some first-person self-knowledge is crucially observational, as when I self-ascribe a present, conscious propositional attitude because, in crucial part, of what it feels like to have it. Of special interest, in the present context, is a third kind of case of first-person knowledge: first-person knowledge of one's own current, conscious propositional attitudes, which is *wholly non-observational*. This is the kind of first-person self-knowledge that I call *purely* first-personal. It is the kind of self-knowledge that I typically have of my own current, conscious beliefs and intentions and that I sometimes have of my own current, conscious desires and emotions. It is also, I contend, the kind of knowledge I have, as agent, of my own future action." (Dunn 2006, p. 38. See 37)

And, similarly, Moran claims: "The authority of the agent does speak from (when he does) as well as the fact that his declaration is made without observation of himself both stem from the fact that the person's own relation to his attitudes and his intentional actions must express the priority of justifying reasons over purely explanatory ones." (Moran 2001, p. 128. See 94–99)

yet, fail to get rid of them insofar as her deliberative answer to questions (Q1)–(Q3) is unable to permeate her dispositions.⁶ To put it another way, she may be able to *report* from a theoretical perspective what desires and beliefs she has and still fail to *express* them:

“So, transparency fails because she cannot learn of this attitude of hers by reflection on the object of that attitude. She can only learn of it in a fully theoretical manner, taking an empirical stance toward herself as a particular psychological subject. We might say that the analysis *reports* on such a belief, but that she does not *express* it, since although she will describe herself as feeling betrayed she will not in her present state affirm the judgment that this person has in fact betrayed her... The distinction between reporting and expressing a state of mind has figured in various philosophical discussions since Wittgenstein.” (Moran 2001, p. 85)

Hence, we can conclude that *an agent has a purely first-person attitude towards herself only when addressing questions (Q1)–(Q3) as deliberative transforms of each other and, therefore, as committed to the transparency condition.* Her answer to such questions involves a *commitment* on her side that *typically permeates* her psychological dispositions to a reasonable degree.⁷ Whenever the previous conditions are met, we can say that the agent’s self-ascriptions *express* a purely first-person point of view. To sum up, Dunn’s and Moran’s approach to *self-knowledge* can be summarized in the four following claims:

⁶ The question “What kind of awareness can have a healing effect in psychoanalytic therapy?” plays a crucial role in Moran’s approach (Moran 2001, pp. 83–94. See Finkelstein 2003, pp. 114–126). The use of the transparency condition to identify the goal of psychoanalytic therapy is one of the ways in which Moran remarkably succeeds at connecting the standard debate about self-knowledge with our ability to lead a meaningful life: “And part of what made the account I develop seem promising to me is that answering this question about how and when Transparency is possible seemed to provide a unified framework for understanding features of self-knowledge that are commonly thought to be fundamental to it, but which are normally discussed in isolation from each other; namely: immediacy and the independence of evidence, the special authority of the first-person, the centrality of the ‘subject-use’ of the pronoun ‘I’, and the importance of ordinary self-knowledge to the rationality of the person.” (Moran 2003, p. 406. See Finkelstein 2003, pp. 153–168; Gardner 2004, p. 250; Moran 2001, p. 107; O’Brien 2003, p. 375; Shoemaker 2003, p. 391) This paper seeks to pursue this endeavor by defending the importance of a certain kind of self-observation for our ability to articulate a genuinely first-person point of view.

⁷ The deliberative model acknowledges the dispositional aspect of our agency. There are a number of dispositions that must be shaped in light of our decisions for an agent to be identified as healthy, but a minimum of permeability is required to be an agent at all and, consequently, to have beliefs, to form intentions or to make decisions. The deliberative model owes us an account, however, of why we are entitled to assume that such a minimum is met and it seems that the most that Dunn and Moran can say in reply is that it follows from the assumption that we are agents with a capacity to deliberate (Burge 1996, pp. 98, 110–111, Moran 2003, p. 405, O’Brien 2003, pp. 378–381; Shoemaker 2003, p. 398), since any appeal to a process of self-observation would allegedly take us away from a purely first-person perspective.

Byrne (2005, pp. 92–98), Gertler (2011, pp. 190–194, 258), and Fernández (2003, p. 365) have defended the dispensability of this transcendental argument and articulated the transparency condition in empiricist terms. I am not sure they have succeeded in their purposes, since it is hard for me to see how the agent’s application of the transparency condition could make sense regardless of some assumptions about her deliberative capacities similar to those posited by Dunn and Moran. This is not, however, an issue I need to take sides on in the present paper.

- (C1) The *deliberative* attitude towards oneself honors the *transparency condition*, that is, it regards questions (Q1)–(Q3) as deliberative transforms of each other. Self-ascriptions on the basis of this attitude are particularly authoritative in the absence of specific *evidence* about oneself and thus *the privilege epistemic authority of self-ascriptions* is accounted for.
- (C2) The deliberative attitude excludes *self-observation* because, otherwise, self-ascriptions of beliefs, decision, and intentions would be based on the *evidence* delivered by a *theoretical, detached attitude* towards oneself that is not specifically first-personal.
- (C3) A purely first-person attitude towards oneself must, by contrast, be *engaged or committed* like that of forming an intention or making a decision.
- (C4) It is a constitutive feature of this engaged, practical attitude that it *typically permeates* the agent's psychological dispositions.

I will, hereafter, refer to this approach to self-knowledge as “the deliberative model.” In the coming sections, I will challenge a crucial assumption in this model, namely: that *evidence and commitment are at odds with each other*, so that any sort of self-observation, insofar as it is concerned with evidence, must essentially be detached or uncommitted as claim (C2) assumes. I will argue, by contrast, that an engaged kind of self-observation is constitutive of an agent's ability to articulate a point of view that could be recognized as purely her own and, therefore, as expressing a genuinely first-person point of view.

For this purpose, I will challenge, in section 2, the claim that an agent's examination of her own dispositions and bodily behavior necessarily belongs to the theoretical attitude as it has been characterized and, in particular, I will defend the *prima facie* normative import of the deliverances of a certain kind of self-observation. As a result, the notion of “expressive self-awareness” will be introduced to characterize this committed kind of self-observation and claim (C2) will be challenged. In section 3, I will argue that questions (Q1) to (Q3) can *only* be construed as deliberative transforms of each other if we assume that some experiences come up with *prima facie* normative authority and, for this purpose, I will rely upon Dunn's (and Blackburn's) own approach to evaluative deliberation. In section 4, I will conclude that part of what the deliberative process might eventually uncover is that the weight attached to one or another reason is not proportional to the situation and must be accounted for in terms of a process of internalization, that is, in terms of the fact that the agent's deliberation has inadvertently been shaped by someone else's views. And I will rely on the notion of expressive self-awareness to elucidate what counts as a genuinely first-person point of view in these circumstances. Expressive self-awareness will thus emerge as essential to our ability to deliberate and, in the end, to articulate a genuinely first-person point of view. It follows that claims (C1), (C3), and (C4) can be defended but only when interpreted in a way that is inconsistent with claim (C2). To develop my argument, I will take advantage of some passages in Carey's novel *Oscar and Lucinda* where the perils of self-deception and the incapacity to rule over one's life are stressed in a rather illuminating way for our purposes.

2 The Wager and a Shaft of Sunlight in a Curtained Room

Oscar felt tempted by a Christmas pudding when he was a child. His relation to the pudding was mediated by his father's conception of it as being sinful. Oscar was not fully aware of this mediation, although some aspects of his behavior, as well as a number of bodily alterations, revealed that he was not completely unaware of it. As the narrator puts it:

“Oscar was not told about the Christmas pudding, but he knew. He did not let himself know that he knew. Yet the knowledge thrust deep into his consciousness. It was a shaft of sunlight in a curtained room. Dust danced in the turbulent air. Nothing would stay still. When Oscar ate his lunch on Christmas Day, his legs ached with excitement. He crossed his ankles and clenched his hands tight around his knife and fork. He strained his ear towards the open kitchen door, but there was nothing to hear except his father breathing through his nose while he ate.” (Carey 1988, pp. 10–11)

In a sense, Oscar knew that his father objected to the pudding, but, as the narrator emphasizes, he didn't know he knew. The fact that Oscar knew is revealed (and determined) by some bodily details (“He crossed his ankles and clenched his hands tight around his knife and fork”) and some attitudes closely associated with certain bodily postures (“He strained his ear towards the open kitchen door, but there was nothing to hear except his father breathing through his nose while he ate”). Oscar's crossed ankles and clenched hands are thus identified in the narrative as part of his response to the situation or, more exactly, as tiny bodily units that express a certain attitude insofar as they form a part of a wider pattern that the agent herself may not have fully grasped yet. Oscar may have missed some of the details mentioned by the narrator about his own body and gestures; moreover, even though he had noticed them, he was in the dark as to what wider pattern they might belong to. Only later, when his father hits him on the back of his head and makes a scolding speech, does Oscar become aware of the pattern at play, namely: he gets to know that he knew that the Christmas pudding was regarded by his father as the fruit of Satan.

The sort of self-knowledge that Oscar thereby acquires is presented as produced by “a shaft of sunlight in a curtained room,” which sounds very close to the metaphor of the mind's eye. We may thereby feel tempted to turn to the Cartesian model of the mind in order to make sense of Oscar's experience. Oscar's mind would thus be depicted as a curtained room and his mind's eye as the shaft of sunlight that illuminates one or another corner of this room. In a similar vein, Dunn and Moran could grant that “a shaft of sunlight” is a sound metaphor to refer to the theoretical attitude and, thus, conclude that its deliverances will only provide a detached sort of self-knowledge.⁸ And this kind of self-knowledge will then fail to express a purely first-person point of view.

⁸ We can safely leave aside here the issue about the privacy of the Cartesian mind. What matters to Dunn and Moran is to oppose the deliberative and the theoretical attitudes, not so much the details as to how the latter could be construed to avoid the traditional concerns with a private mind. In any event, the engaged sort of self-observation I intend to sketch in this paper can hardly be conceived of as private, given that its deliverances are claimed to form a pattern of intelligibility with the situation the agent faces.

There are some elements in the novel that favor this interpretation. There is, for instance, a tendency to leave the life of the characters in the novel unaltered by whatever their respective shafts of sunlight may deliver. The contents that such shafts bring out are often presented as idle (and, therefore, as apparently dispossessed of any *prima facie* normative authority), since they hardly ever lead the agent to modify her behavior or her attitudes correspondingly. In general, the novel seems to convey a sense of dismay and fatality in an agent's getting to know her own psychological condition. The fundamental attitudes of the main characters in the novel, namely, Oscar and Lucinda, do not change throughout their lives. They seem to be seized by their fates. They struggle to lead their lives in a meaningful way but they seem to be trapped within the demands of their puritanical upbringing and the constraints of the rigorous society they live in. We may thus feel tempted to interpret the inertness of their respective mind's eye as a constitutive feature of this sort of experience, which is what the deliberative model suggests.

We could instead regard this inertness as an idiosyncratic feature of Oscar's and Lucinda's character associated with their specific incapacity to lead a meaningful life (Corbí 2012, pp. 137-155). In the coming paragraphs, I will try to motivate the second option; more specifically, I will argue that those experiences that shafts of sunlight burst into consciousness do typically emerge as endowed with *prima facie normative authority*, whereas it is their inertness—that is, their inability to shape or permeate the agent's lives what must be accounted of in terms of some sort of interference or disruption.

Let us consider, for this purpose, a particular wager in which Oscar's and Lucinda's lives got entangled. At a late stage in the novel, they bet at each other's fortune on whether Oscar would be able to transport and install a Crystal Church in a remote settlement at Bellingham where Reverend Dennis Hasset had been exiled. The project was close to madness given the difficulties of the design, the pointlessness of a transparent church in a sun-beaten location, and, finally, the hurdles of transportation across the island due to Oscar's dread of the ocean. The apparent vanity of the project emerged clearly into Lucinda's mind when confronting a third party's view, namely, that of Mr d'Abbs':

“As Mr d'Abbs spoke and as Lucinda looked at this tawdry church she began to suffer a tight, airless feeling in her chest. The fact that the object of their bet was now made to appear at once so vain and mediocre and that it was, in any case, impossible to build, conspired to act as a catalyst in Lucinda's soul, to make a focus for all the vague unease she harboured about the bet, and fearful thoughts which she had hitherto managed to keep submerged, now bubbled up like marsh gas and burst, malodorous, in the very forefront of her conscious mind. The tight band across her chest was not an unfamiliar feeling. It normally came on her after a night spent at the gaming tables. It was a panic produced by the fear of throwing away her fortune. She pressed her forearms against her abdomen. She looked to Oscar, wishing only that he would dispel her panic with a smile.” (Carey 1988, p. 416)

Mr. d'Abbs words acted as a catalyst of a number of experiences (including “the vague unease she harboured about the bet” and “the fearful thoughts which she had

hitherto managed to keep submerged”) that Lucinda could easily have noticed but only these words brought “to the very forefront of her conscious mind” as elements of a unified pattern. This new experience could thus be described as the dawning of an aspect whose emergence had largely been suppressed. The important question for our purposes now is whether the dawning of this aspect, the formation of this gestalt, comes up with *prima facie normative authority* or, in other words, whether *they constitutively call for a certain response*. In case of a positive answer to this question, it will not be the existence of a response that must be accounted, but its absence. And, to this end, we ought to mention one or another element that may have blocked or canceled the agent’s response, namely: a set of reasons that may have counteracted or undermined the normative authority of the new aspect perceived or else a process of suppression in which the agent may have indulged.

Lucinda’s experience in the passage above seems to favor a positive answer to the previous question. What the shaft of sunlight brings out does have an impact of its own, and this is a recurrent feature in both Lucinda’s and Oscar’s experiences as they are described in the novel. Some tiny bodily changes and gestures (“pressing her forearms against her abdomen”) are often highlighted to portray their attitude towards the situation, the way they respond to it. Such bodily changes could hardly be expressive of a certain attitude, however, if they were not causally linked to some inner experiences or thoughts. But it is not only causal efficacy that is needed; those bodily alterations must also form a part of a pattern of response that is *proportional* to the situation to a minimal degree for an expressive link to be at all intelligibly identified. This notion of proportionality is not foreign, however, to the deliberative model, since the deliberative attitude must honor the transparency condition and, consequently, it is assumed that an agent’s decisions and commitments must be proportional to the situation to a certain degree. The question then arises as to whether the sort of bodily alterations we have so far identified do constitute part of a proportional response to the situation Lucinda has so painfully become aware of.

It is clear that the tiny bodily alterations mentioned by the narrator can hardly constitute such a proportionate response and, yet, they emerge as part of a response. It is suggested in the passage, however, that some elements of *suppression* may lie behind the fragmentation in her response. In fact, she suppressed (“fearful thoughts which she had hitherto managed to keep submerged”) some aspects of her own response to the Crystal Church project, so that it might not appear to her conscious mind as vain and mediocre; in other words, she suppressed this view of the project in order to preserve her commitment to it. So, it seems that, when an agent’s evaluative attitude or view thrusts into her consciousness, it permeates her dispositions despite her efforts to the contrary and in the direction of the specific normative import they come up with. Hence, insofar as shafts of sunlight are interpreted as an apt metaphor for the experience of self-observation, some reason has been provided to deny that the deliverances of self-observation should constitutively be construed as essentially deprived of *prima facie normative authority*, that is, as intrinsically alien to the sort of proportionality in the agent’s response that the deliberative attitude demands. We must, instead, allow for attitudes or views to be thrust into consciousness as endowed with *prima facie normative authority*, that is, with a certain normative import that is independent of its instrumental value with regard to some other dispositions or attitude that the agent might actually have.

It includes a certain normative import but only *prima facie*, since the agent may eventually *estrangle or disown* those evaluative attitudes or views insofar as they may be at odds (and not only instrumentally) with some other concerns or experiences of hers. This maneuver of estrangement seeks to cancel the normative import of a certain evaluative attitude or view, so that it may no longer shape the agent's life. This is not to deny that success in this endeavor should often count as a remarkable achievement, since some attitudes are so deeply ingrained that one can hardly disown them, at least in the strong sense of depriving them of their capacity to shape one's life, including not only one's psychological dispositions but the way one deliberates. We can thus say that an agent becomes *expressively self-aware* of a certain evaluative attitude or view of hers when it emerges into her consciousness in such a way that she feels *prima facie committed to it* (that is, as endowed with *prima facie normative authority*), even though she may eventually try to *estrangle or disown* it inasmuch as she may regard this attitude or view as *calling for a response* that, in light of her other concerns and attitudes, is not *proportional* to the situation.

Let's now see how this notion of expressive self-awareness may help us to address a crucial issue regarding the deliberative attitude, namely, what resources might contribute to a proper answer to questions (Q1) and (Q3) so that they could coherently be regarded as deliberative transforms of each other. Moran hardly ever addresses this question in any detail, but we can rely for this purpose on Dunn's elaborate account of practical deliberation. I will argue, however, that Dunn's account cannot coherently exclude the engaged kind of self-observation I have just sketched from an agent's practical deliberation and, as a result, I will conclude that his view on practical deliberation conflicts with his understanding of the deliberative attitude as wholly non-observational.

3 Deliberative Transforms and Shifts of Focus

Dunn, following up on Simon Blackburn, explores Christine Korsgaard's invitation to distance oneself from one's own desires and concerns and to examine them in light of an independent, moral standpoint. As Blackburn points out, it makes sense to carry on such an assessment independently of any *particular* desire an agent might have, but it can't plausibly be done regardless of *all* her desires and concerns.⁹ How are we then supposed to proceed? Blackburn and Dunn argue that the focus of practical deliberation is not so much one's desires and concerns but those features of the world that become salient as a result of them:

“Korsgaard's account is striking in the way it pictures the deliberative setting: when we deliberate about what to believe or to do, our topics of consideration are our own psychological states - our perceptions or our desires. Now, this is an odd idea. It doesn't seem right at all. Even if some of our evidence is about how the world presents to us, most of it is about how the world is; and, even if we

⁹ Thus, Blackburn claims: "... In the sense in which it is right, it means only that one can stand back from a *particular* desire or impulse, and accept or reject its pressure on one. Certainly we can do this, in the light of other desires and concerns. What is not thereby given is that we can do it from a standpoint independent of any desire or concern: independent of a desire for our own good, or for the happiness of humanity; or respect for this or that, or the myriad other passions that make up our individual profiles of concern and care." (Blackburn 1998, p. 252)

sometimes explicitly take account of our desires when deciding what to do, very often we focus on features of the world around us that matter to us in some way.” (Dunn 2006, p. 84)¹⁰

So, Korsgaard seems to incur in what Blackburn and Dunn call “the objectifying mistake,” namely, the assumption that deliberation on what to do should mainly focus on the plurality of one’s desires in order to decide which to endorse. But how does this shift from one’s desires to the world helps us to address the issue as to how one must deliberate regarding questions (Q1) to (Q3)?

This shift of focus is explicitly associated in Dunn’s account with the distinction between figure and background which, in turn, provides the framework within which we can claim that questions (Q1) to (Q3) are deliberative transforms of each other and, therefore, that they meet the transparency condition. More specifically, question (Q2) focuses on certain features of the world while the agent’s desires and concerns lie in the background, but question (Q3) shifts the focus to the agent’s psychological condition and the relevant features of the world are left in the background. So, the question arises as to how the features of the world are to be individuated for them to favor a certain answer to question (Q2) and, in particular, whether they could be individuated from a theoretical perspective given that they must also contribute to question (Q3) in a way that is not merely instrumental. It seems that we cannot exclusively rely on the deliverances of this sort of perspective because they are deprived of any *prima facie* evaluative import that might justify a certain answer to question (Q2) and, therefore, to question (Q3). Hence, it seems that if some features of the world must contribute to an answer to question (Q2) in such a way that they non-instrumentally contribute to an answer to question (Q3), they must be granted a certain *prima facie* normative import. We must then make room for a kind of exploration of the world—i.e., a way of gathering evidence—that departs from the theoretical perspective and such that its deliverances are endowed with *prima facie* normative authority, that is, call for a certain response on the agent’s side.¹¹ Once we allow for this kind of exploration of the world, I do not see how Dunn could exclude a similar kind of

¹⁰ In a similar vein: “... It surely is a mistake to suppose that our desires ‘fill the foreground’ of practical deliberation. For one thing, the phenomenology of reflection doesn’t lend any support to the idea that we are self-preoccupied like this as practical deliberators. The situation is at least very often as Blackburn depicts it: we consider this and that, where this and that are the various features that weigh with us as we try to decide what to do.” (Dunn 2006, p. 89)

¹¹ This line of reasoning implies that there are psychological states with a dual direction of fit, that is, a world-to-mind and mind-to-world direction of it. The intelligibility of such states is explicitly defended by Dunn regarding evaluative beliefs: “One and the same state—the belief that p ought to be—has a mind-to-world direction of fit with respect to the content that p ought to be (marking out its states as a belief) and a world-to-mind direction of fit with respect to the embedded content that p (marking out its status as, or as a state that involves, the desire that p).” (Dunn 2006, p. 17)

Hence, in assuming the existence of this sort of psychological state, I am not going beyond what Dunn’s deliberative model concedes. Moran is less explicit in this respect. Still, some commentators have point out some difficulties in his approach that might be solved by allowing for attitudes with a dual direction of fit. Gardner (2004, p. 262) is puzzled about how a fact could have practical significance and still be recognized as a fact from a realist perspective, whereas O’Brien (2003, pp. 366–7, 381–2) straightforwardly suggests that some sort of non-conceptual practical awareness is required to explain “... how agency gives us knowledge” (O’Brien 2003, p. 367) and it seems that practical awareness could not but have a dual direction of fit.

attitude towards on one's own attitudes or psychological condition, since a shift of focus from the object to one's psychological dispositions, and vice versa, is central to the notion of deliberative transform. This is, though, the kind of self-exploration that the notion of expressive self-awareness is meant to apprehend.

In light of this notion, an agent's traits of character could contribute to her practical deliberation in a non-instrumental manner, since they will typically emerge as endowed with normative authority, even though this authority will only be *prima facie* insofar the agent may eventually disown or estrange any particular trait of her character.¹² It follows from the previous discussion that an agent could not determine her traits of character by focusing on her inner experience alone, but on how she responds to a number of situations and, therefore, on the pattern of intelligibility formed by those situations and the way the agent actually responds to them. We can thus conclude that the deliberative attitude (and, therefore, a construal of questions (Q1) to (Q3) as deliberative transforms of each other) implies an examination not only of the situation but of the agent's character as well, for only by these means could the proportionality of a certain commitment or decision be determined. To this effect, it is important to highlight that an agent's character can hardly be construed as an additional feature of the situation, since an agent's commitment to one or another aspect of her character, to the normative authority that comes up with it, is subject to revision in the process itself of deliberation whereas the different features of the situation must be approached as given, as constraints, in view of which to deliberate.

In any event, the role I attach to an agent's traits of character in her practical deliberation squares quite nicely with Blackburn's and Dunn's account of practical deliberation and their emphasis on the shift of focus between the situation and the agent's psychological condition. It conflicts, though, with Dunn's attempt to individuate a wholly non-observational kind of self-knowledge as constitutive of the deliberative attitude. In section 4, the notion of expressive self-awareness will be relevant to the kind of deliberation in light of which a certain view, however internalized, can legitimately be recognized as genuinely one's own. Thus, claims (C1), (C3), and (C4) will be confirmed but only insofar as they are interpreted in a way that is inconsistent with claim (C2) and, therefore, with the idea that a genuinely first-person point of view must be wholly non-observational.

4 The Fruit of Satan: "They Would Make Me into the Creature They Perceived."

Oscar tasted the Christmas pudding. He treasured it in his mouth. He had a view about it. It tastes lovely:

"Oscar took the spoon and ate, standing up.

¹² In section 4, I will emphasize how the agent may challenge some traits of her character and, yet, fail to estrange it insofar as her response to some situations may still be shaped by it.

He could never have imagined such a lovely taste. He let it break apart, treasuring it inside his mouth.” (Carey 1988, p. 12)

He was able to keep his view despite his father’s scolding. Theophilus treated it as poison, but his son internally rejected his views, “it did not taste like the fruit of Satan.” Oscar opposes his own experience of the pudding to his father’s condemnation and he manages to resist it. His father could confine his son to a room and prevent him from eating the pudding, but he was unable to change his son’s mind. There is a sense in which Oscar had his own view about the pudding.

“Oscar had never been hit before. He could not bear it.

His father made a speech. Oscar did not believe it.

His father said the pudding was the fruit of Satan.

But Oscar had tasted the pudding. It did not taste like the fruit of Satan.” (Carey 1988, p. 12)

There is in Oscar’s experience an expressive unity of body and soul. It is because the pudding tastes lovely that it can’t be the fruit of Satan, no matter what his father or some other authoritative voice might claim on this matter. There is no mismatch between Oscar’s views and his dispositions simply because, at this stage, he trusts his dispositions as a criterion of the good. It sounds then that his views are sensitive or permeable to his experiences rather than the other way round (Corbí 2010, 2011, 2012, pp. 156-186). How long will Oscar be able to keep the autonomy of his views, though?

Not for long. His contact with the pleasurable aspects of life is severely mediated by his father’s puritanical outlook. When he goes to the theater he feels divided. There are two minds in a single body: “Oscar hesitated. He saw the theatre with two sets of eyes, one his own, but one his father’s. The second set saw the theatre steeped in sin” (Carey 1988, p. 111). The experience of himself as divided allows Oscar to preserve a clear sense of his own identity. So, it seems that an agent can detach herself from those thoughts and attitudes that derive from a view that she perceives as alien and identify herself with those that emerge from an entirely different perspective. This is the way the plight of a psychoanalytic patient is typically characterized. The problem is that the agent’s estrangement from a certain internalized view does not cancel the capacity of the latter to permeate her dispositions, so that the agent will still be unable to act and experience her life the way she wants to, that is, from a viewpoint she identifies herself with. The inability to implement one’s own view is thereby explained in terms of a divided self.

There are, however, some situations that depart from this standard psychoanalytic case inasmuch as the agent may not realize how a particular decision, attitude, or action do not really express her own views, goals, and purposes, but just those of a third party, typically that of her parents’. For instance, Lucinda moved to Sidney just after her parents died and she bought a

Glassworks factory with the fortune she had inherited. When Lucinda arrives in Sidney, she feels uncertain about her future, although the particular way in which she experiences this uncertainty is not, according to the narrator, genuinely personal but the result of enacting someone else's view, namely, that of Sol Myer's:

“The river journey was picturesque, with so many pretty farms along its banks. Lucinda could not look at them without feeling angry. She looked straight ahead, shivering. It was cold, of course, but not only cold that caused this agitation. There was a jitteriness, a sort of stage fright about her future which was not totally unpleasant. She dramatized herself. And even while she felt real pain, real grief, real loneliness, she looked at herself from what she imagined was Sol Myer's perspective, and then she was a heroine at the beginning of an adventure.” (Carey 1988, pp. 130–131)

How is it that Lucinda happens to experience her future from Sol Myer's perspective? It seems that it should be the product of some process of *internalization*. Yet, in contrast with my previous remarks about Oscar, Lucinda is rather unaware of this process and, therefore, what she regards as her own view on a certain matter turns out to be Sol Meyer's without further reflection on Lucinda's side. And such alien view has the power to permeate both her dispositions and her own experience. Yet, the ability of someone else's view to shape Lucinda's experience is not confined to those cases in which she may be unaware of its influence. For instance, she was aware at some point of the impact that her employees' views have on herself, but was unable to prevent their look from transforming her own experience and attitudes:

“By the way they looked at me, by their perception of me, they would make me into the creature they perceived. I would feel myself becoming a lesser thing. It is the power of men.”

“But I am a man.”

“No,” she said, too impatient to let him develop his argument. ‘Of *men*, men in a group, men in their certainty, men on a street corner, or in a hall. It is like a voodoo. Do you know voodoo?’” (Carey 1988, p. 146)

So, it seems that the way men in a group look at Lucinda has an impact on her to the point of making of her the kind of creature they perceive. The only way she has to forsake this effect is to avert their gaze. It follows that the look of these men is not experienced as at all inert. *Their look expresses a view about Lucinda that shapes her identity and, therefore, the way she will face any given situation in the future.* This may seem mysterious, like a voodoo, but it is similar to what happens when Oscar's internalizes his father's view to the effect that he looks at the stage in the theater with two eyes, or when Lucinda experiences her arrival in Sidney from Sol Myer's view. All these experiences pose a problem as to the conditions in which an agent could recognize a point of view as genuinely her own. In fact, I will argue that the deliberative model cannot coherently describe those conditions while the notion of expressive self-awareness provides a proper basis for this description.

What the deliberative model regards as purely first-personal is the agent's capacity to make decisions, to form intentions or to acquire beliefs in ways that respect the transparency condition and such that they typically permeate her dispositions. But the previous passages suggest that neither transparency nor permeability as such guarantee that the agent's ability to express a first-person point of view, since someone else's view can shape her deliberation and have a rather similar impact on the agent's dispositions. What the agent regards as her views can easily be the product of an unnoticed process of internalization. Moran is certainly eager to distinguish an agent's control over her own attitudes as a result of adopting a deliberative attitude from the sort of control that she can have over someone else's mind:

“Beliefs and other attitudes, on the other hand, are stances of the person to which the demand for justification is internal. And the demand for justification internal to the attitudes involves a sense of agency and authority that is fundamentally different from the various forms of direction or control one may be able to exercise over some mind or another.” (Moran 2001, p. 114)

This contrast is central to the deliberative model, as we see. My worry is, though, whether the deliberative model is in a position to distinguish these two sorts of control. It is true that permeability involves a sort of directness that contrasts with the indirect way in which an agent may be cheered up by drinking a beer or someone may try to make you angry by insulting your brother. But, in cases of internalization, the impact of someone else's views is not indirect in this trivial sense. In those cases, other people's views seem to have a rather more direct impact on the agent's dispositions because they inadvertently shape the way the agent faces the situation and, therefore, how she identifies the salient features in the situation and how they are to be pondered; in other words, other people's views may inadvertently shape how the agent deliberates. The deliberative model must then provide a more fine-grained understanding of what “direct” means if we are to discard someone else's control over an agent's attitudes as indirect. In the coming paragraphs, I will suggest how the notion of expressive self-awareness allow us to meet this demand, that is, to articulate a way of being in control of one's life “that is fundamentally different from the various forms of direction or control one may be able to exercise over some mind or another.”

Consider again Oscar's taste of the Christmas pudding. He is convinced that it is not the fruit of Satan because it tastes lovely. Oscar's faithfulness to his body provides a basis upon which he affirms his views against his father's. He trusts his bodily experiences and acknowledges their authority to determine his reasons to act. The situation becomes more confusing, though, as we depart from this primeval experience. In the theater, Oscar is no longer confronting his father's view with his bodily experience in such a neat manner, even the pleasure that the play provides contributes to its value from one point of view and to its disvalue from the other. The more entangled one's experience becomes, the harder it is to identify one's bodily experiences and their precise *prima facie* normative import. The narrator is often engaged in the process of disentangling the characters' experiences by focusing on tiny bodily alterations, thoughts, feelings, and responses. But all this is conducted from a third-person perspective. The characters themselves only engage in such a task quite sporadically and even then in a rather erratic manner, as happens when Lucinda

becomes aware of the impact of the look of men in a group. An agent could, in any event, turn to her own tiny bodily experiences and seek to ascertain their meaning in order to articulate a voice that could be identified as genuinely hers. Such a voice could not be exclusively hers because any agent's experience, including her bodily experiences, has already been shaped by many authoritative voices. But, at least, an agent can try to articulate a life that *makes some sense to her* given the way she has actually been shaped. We can also say that an agent's response to a situation will make sense to her if it is *proportional* both to the situation and to her own character. And we can regard this experience of making sense as fundamental for the agent's ability to articulate a view or a reason that is genuinely hers, regardless of whether it might have been previously internalized. Expressive self-awareness is the sort of exercise that may lead the agent to shape a response that could eventually meet this condition, for this kind of self-awareness focuses on one's psychological condition as part of one's response to a situation and, therefore, on the pattern formed by the situation and one's response to it.

Needless to say, the experience of making sense allow for a significant degree of uncertainty and indeterminacy as to what views should count as genuinely one's own. And there is no epistemic privilege of the first-person perspective as to whether a certain project or engagement really contributes to making sense of one's life or, in other words, as to whether a certain self-adscription is genuinely expressive of one's own views. This does not undermine, however, the agent's epistemic privilege regarding what views he explicitly endorses or what intentions she has just formed, no matter whether they are genuinely hers or no. I have no room, however, to explore whether the deliberative attitude as I have elucidated it can still account for these epistemic privileges along the lines suggested by Dunn and Moran, or an alternative approach should be sketched. This issue should be left for another occasion. We can conclude, in any event, that claims (C1), (C3), and (C4) can be defended but only when interpreted in a way that is inconsistent with claim (C2) and, thereby, some reason has been provided to reject a central aspect of the deliberative model of self-knowledge as presented by Dunn and Moran.

Acknowledgements I am indebted to my colleagues in the Phronesis Weekly Seminar for their generous discussion of an early draft of this paper. I have also benefited from the questions and remarks by the participants in VII SEFA Conference (Madrid, September 11-14, 2013), Workshop on Agency, Responsibility and Self-Deception (Bogotá, November 21-22, 2013) and Workshop on Self-Knowledge, Expression and Transparency (Murcia, March 17-18, 2014). I am finally pleased to acknowledge that research for this paper has been funded by Spanish Ministry of Economy and Competitivity (CSD2009-00056, FFI2013-47948-P, FFI2014-55256-REDT).

References

- Blackburn, S. (1998). *Ruling passions*. Oxford: Oxford University Press.
- Boyle, M. (2009). Two kinds of self-knowledge. *Philosophy and Phenomenological Research*, 78, 133–164.
- Burge, T. (1996). Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society*, 96, 91–116.
- Byrne, A. (2005). Introspection. *Philosophical Topics*, 33, 79–114.
- Carey, P. (1988). *Oscar and Lucinda*. London: Faber and Faber.
- Corbí, J. E. (2010). First-person authority and self-knowledge as an achievement. *European Journal of Philosophy*, 18, 325–326.

- Corbí, J. E. (2011). Observation, character, and a purely first-person point of view. *Acta Analytica*, 26, 311–28.
- Corbí, J. E. (2012). *Morality, self-knowledge and human suffering: an essay on the loss of confidence in the world*. New York: Routledge.
- Dunn, R. (2006). *Values and the reflective point of view: on expressivism, self-knowledge and agency*. Aldershot: Ashgate.
- Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.
- Fernández, J. (2003). Privileged access naturalized. *The Philosophical Quarterly*, 53, 352–372.
- Finkelstein, D. (2003). *Expression and the inner*. Cambridge: Harvard University Press.
- Gardner, S. (2004). Critical notice of Richard Moran, authority and estrangement: an essay on self-knowledge. *The Philosophical Review*, 113, 249–267.
- Gertler, B. (2011). *Self-knowledge*. New York: Routledge.
- Moran, R. (2001). *Authority and estrangement: an essay on self-knowledge*. Princeton: Princeton University Press.
- Moran, R. (2003). Responses to O'Brien and Shoemaker. *European Journal of Philosophy*, 11, 402–419.
- O'Brien, L. (2003). Moran on agency and self-knowledge. *European Journal of Philosophy*, 11, 375–390.
- Sartre, J. P. (2003). *Being and nothingness*. New York: Routledge.
- Shoemaker, S. (2003). Moran on self-knowledge. *European Journal of Philosophy*, 11, 391–401.
- Zimmerman, A. (2008). Self-knowledge: rationalism vs. empiricism. *Philosophy Compass*, 3, 325–352.