

FIRST PERSON AUTHORITY AND KNOWLEDGE OF ONE'S OWN ACTIONS

MARTIN F. FRICKE

Instituto de Investigaciones Filosóficas
y Centro Peninsular en Humanidades y Ciencias Sociales
Universidad Nacional Autónoma de México
martin_fricke@yahoo.co.uk

SUMMARY: What is the relation between first person authority and knowledge of one's own actions? On one view, it is because we know the reasons for which we act that we know what we do and, analogously, it is because we know the reasons for which we avow a belief that we know what we believe. Carlos Moya (2006) attributes some such theory to Richard Moran (2001) and criticises it on the grounds of circularity. In this paper, I examine the view attributed to Moran. I rebut the charge of circularity, but also reject the theory as an adequate interpretation of Moran.

KEY WORDS: self-knowledge, transparency, belief, Richard Moran, Carlos Moya

RESUMEN: ¿Qué relación existe entre la autoridad de la primera persona y el conocimiento de las propias acciones? Una posibilidad es que gracias al conocimiento de las razones que tenemos para actuar sabemos qué es lo que hacemos y, análogamente, gracias al conocimiento de las razones que tenemos para admitir [*avow*] una creencia sabemos qué es lo que creemos. Carlos Moya (2006) atribuye una teoría de este tipo a Richard Moran (2001) y la critica por ser circular. En este trabajo examino la teoría atribuida a Moran, refuto la idea de circularidad, pero también rechazo la teoría como una interpretación adecuada de Moran.

PALABRAS CLAVE: autoconocimiento, transparencia, creencia, Richard Moran, Carlos Moya

In *Crítica*, vol. 38, no. 114 (2006), Carlos Moya ascribes a theory of self-knowledge to Richard Moran according to which knowledge of one's own actions explains authoritative knowledge of one's own beliefs. This is because, on this interpretation of Moran's theory, knowing of one's own beliefs is a special case of knowledge of one's own actions. Moya criticises this account of self-knowledge on the grounds of circularity. In this discussion note, I examine the relation between the two types of knowledge in Moran's theory. I shall try to show that Moran's reference to knowledge of one's own actions is not an attempt to give the finer detail of how we acquire self-knowledge with respect to beliefs. Rather, it is an illustrative analogy. But perhaps Moya is right in that Moran's account of knowledge of one's own beliefs is incomplete. If it is, then, so I shall suggest, his account of knowledge of one's own actions is as well.¹

¹ Many of Moran's readers think that his account is important (e.g. Heal 2004,

Someone who sincerely affirms that she believes that P rarely makes a false utterance. It might be that not P; but it is unlikely, some would say impossible, that she does not *believe* that P. The self-knowledge expressed in “I believe that P” is not only less prone to error but also more immediate than knowledge about the beliefs of other people. To find out whether someone *else* believes that P I have to observe the other person and, perhaps, draw more or less implicit inferences based on these observations. In order to know that *I* believe that P, I do not seem to have to make such observations or draw such inferences. How should such first person authority be accounted for?

Moran offers an explanation of both aspects of first person authority which is based on the idea that the question “Do I believe that P?” is *transparent* to the question “Is it true that P?” The idea of transparency, as Moran understands it, is expressed in the following well-known quote from Gareth Evans:

If someone asks me “Do you think there is going to be a third world war?”, I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering the question “Will there be a third world war?” I get myself in a position to answer the question whether I believe that *p* by putting into operation whatever procedure I have for answering the question whether *p*. (Evans 1982, p. 225)

Here, the question about my belief about a third world war is “transparent” to a question about the war itself. “A first-person present-tense question about one’s belief is answered by reference to (or consideration of) the same reasons that would justify an answer to the corresponding question about the world” (Moran 2001, p. 62).

Self-knowledge acquired in this way might be expected to enjoy first person authority in both of the senses described above. Going from a positive answer to “Is it true that P?” to a positive answer to “Do I believe that P?” seems easier than to observe a person and then draw inferences about her beliefs. This might explain why we are less prone to error in such self-ascriptions. Likewise, it seems that the procedure is at least ‘more direct’ than those we have for

Shoemaker 2003, Byrne 2005, O’Brien 2003); but there is surprising disagreement about what it consists in. Even a colleague at his own department is unable to say plainly and simply what his theory is and only offers his thoughts as to how Moran “is best understood”, while “not claim[ing] that Moran himself understood his idea in exactly this way” (Boyle 2011, p. 235).

ascribing beliefs to other persons, since it is not dependent on truthful observation. So it is easy to see why the transparency described by Evans is often seen as central to an explanation of first person authority.

However, Moran's central question does not seem to be how transparency contributes to authority. Rather, he wants to know how such transparency is *possible*. How can it be that we answer a question about someone's mind (what do I believe?) by considering what is the case in the world? This question about the possibility of transparency can be understood in at least two different ways (cf. Shoemaker 2003, p. 400). First, we can ask as to the metaphysics involved in transparency. What mental mechanisms (functional roles, cognitive dynamics, etc.) allow that a consideration of questions about the world produce knowledge of one's own mind? Second, we can understand the question as an epistemological one. What epistemic *right* does a subject have to consider questions about the world when answering questions about her beliefs? What *legitimises* this procedure? Moran, clearly, is only interested in this second sense of the how-possible question.

His answer, in its most concise form, can be found in a discussion with Lucy O'Brien and Sydney Shoemaker: "[I]t is only because I assume that what I actually believe about X can be determined, made true by, my reflection on X itself, that I have the right to answer a question about my belief in a way that respects the Transparency Condition" (Moran 2003, p. 406). Moran adds that the assumption in question can be interpreted as a Kantian "Transcendental Assumption of Rational Thought" (Moran 2003, p. 406). If my reflection (or deliberation) about X did not determine what it is that I believe about X, then why should I engage in such reflection? The purpose of rational thought is the acquisition of beliefs. It only makes sense to undertake such reflection or thought if this determines my beliefs. For this reason, Moran talks of an assumption of rational thought. If the assumption is true, then I have the epistemic right to self-ascribe beliefs on the basis of my reflection about the world. So Moran's answer to the question of how transparency is possible is this: We have an epistemic right² to transparency in as much as we

² Moran also claims that I normally have an *obligation* to self-ascribe my beliefs only on the basis of reasoning about the world (cf. Moran 2003, pp. 415 ff.). But this obligation does not seem to follow from the Transcendental Assumption of Rational Thought. (The purpose of my thought about the world is to form true beliefs. The purpose is *not* to make correct self-ascriptions of belief. So someone who rarely makes self-ascriptions of belief or does not generally base them on Evansian

are capable of rational thought. Transparency is legitimate in so far as we are rational thinkers.³

Moran often mentions the knowledge we have of our own actions, when he tries to explain his theory of the self-ascription of beliefs: “A central example for me is that of making up one’s mind about what to do, where the answer arrived at expresses first-person knowledge of my own future action” (Moran 2003, p. 411). Why is the knowledge we have of our own actions a “central example” for Moran? How does it contribute to our understanding of the authority we have in self-ascriptions of belief? There are passages in Moran’s book that suggest that knowledge of one’s own actions can provide special clues to understanding the knowledge we have of our own beliefs. A crudely simplified argument might go as follows:

- (i) I know what I am doing.
- (ii) I make my beliefs.
- (iii) Therefore, I know what beliefs I have.

Both premises of this argument are questionable. Premise (i) does not seem to be true of unintentional actions. What makes such actions

transparency does not thereby seem to violate the Transcendental Assumption. It would still be true that what she believes about X is determined by her reflection about X. (She would just not be very good at knowing what she believes.) Moran’s discussion suggests that on his view the obligation arises because cooperation with others is impossible if they cannot take our self-ascriptions of beliefs and other intentional states as the result of our reasoning (2003, pp. 415 ff.).

³There is a slight tension between this formulation and the above quote (Moran 2003, p. 406). On my reformulation, it is the fact *that I am a rational thinker* that legitimises transparency. In the quote from Moran, it is the fact *that I assume* that I am a rational thinker that legitimises transparency. While it might be true that the assumption is necessary for any thought (this is what makes it “transcendental”), it is not clear to me that the mere assumption is sufficient to confer epistemic legitimacy on transparency. It seems to me that the assumption must also be true. (The assumption that the moon is made of cheese confers very little, if any, legitimacy on my belief that it is edible.) We can probably say that Moran is aware of this further requirement (that the assumption be true) in so far as he recognises cases of compromised rationality where we make the assumption but are in fact incapable of actually forming our beliefs according to our deliberation and therefore do not enjoy first person authority. (I am grateful to a referee for pointing this out.) Is the truth of the assumption sufficient to legitimise transparency? On some accounts of epistemic legitimacy, it might be necessary to give a full metaphysical account of how transparency comes about in order to answer the question about its legitimacy. On such accounts, the above distinction between a metaphysical and an epistemological answer to the question of how transparency is possible collapses.

unintentional seems to be precisely that I do *not* know what I am doing. But perhaps premise (i) is true if the action in question is intentional. Still, what reasons can be given for holding it to be true in this case and how is it possible that it be true? But premise (ii) is even more problematic. In what sense do we make our beliefs? If premise (i) is true only of intentional actions, then, for the argument to be valid, premise (ii) must also describe an intentional action. But is it true that we make our beliefs intentionally? It seems that most beliefs, even those we know about authoritatively, come to us rather unintentionally. We cannot help but having them.

Carlos Moya suggests that a more complete argument of the above form can be found in Moran's book (cf. Moran 2001, pp. 125 ff.) in the context of a discussion of Elizabeth Anscombe's theory of intentional actions (cf. Anscombe 1963). There Moran writes:

The description under which an action is intentional gives the agent's primary reason in so acting, and the agent knows this description in knowing his primary reason. This description is known by him because it is the description under which he conceives of it in his practical reasoning. It is the description under which the action is seen as choiceworthy by him, as aiming at some good to be achieved. (Moran 2001, p. 126)

According to Moya, Moran here explains why intentional actions are known by the agent, as suggested by Anscombe. The idea is that intentional actions are the result of a deliberation, a process of practical reasoning. In such practical reasoning we consider the (primary) reasons for our actions and these reasons contain descriptions of the actions for which they are reasons. If we now add that avowing and endorsing a belief is likewise an intentional action based on prior reasoning (deliberation), then it seems to follow that we must know the beliefs so avowed and endorsed:

Avowing and endorsing a belief are intentional actions themselves, which, following Anscombe, I would know without observation (or discovery). [...] Now, applying the line of explanation that Moran suggests concerning non-observational knowledge of our intentional actions to the case of beliefs formed and avowed as a result of deliberation, knowledge of these beliefs would owe to the fact that the corresponding belief, as it is avowed, would appear in the agent's deliberation, maybe as entertained or considered in the process of weighing up reasons for and against it. (Moya 2006, p. 11)

We might analyse this argument as follows:

- (1) Intentional actions are based on a process of practical reasoning (deliberation).
- (2) In such practical reasoning we consider the primary reason of our action.
- (3) The primary reason of an action includes a description of the action.
- (4) Therefore, we know our own intentional actions.
- (5) Avowing and endorsing a belief is an intentional action.
- (6) Therefore, avowing and endorsing a belief is based on a process of reasoning (deliberation).
- (7) In such reasoning we consider the primary reason for avowing and endorsing the belief.
- (8) This primary reason includes a description of the belief to be avowed and endorsed.
- (9) Therefore we know the beliefs we avow and endorse.

According to this argument, knowledge of one's own beliefs is a special case of knowing one's own action. It is derived from the fact that in deliberating about whether or not to perform some action, namely that of avowing and endorsing a belief, we consider reasons that contain descriptions of this belief.

In what follows, I shall consider an objection of Moya's to this argument. I will try to show that the objection is misguided. I will also claim that the argument fails as a reconstruction of Moran's theory and will suggest a different interpretation.

Moya's main objection to the reconstructed argument is that it is circular. The primary reason of an action—at least if understood in a Davidsonian way—consists of a belief and a pro-attitude. So knowing the primary reason of an action involves knowing a belief of oneself, namely the belief that, together with some pro-attitude, constitutes the primary reason for our action. Moya claims that this means that, according to this theory, knowledge of one's own action *presupposes* knowledge of some belief of oneself. It therefore makes no sense to suppose that knowledge of one's own actions can explain, as a model, knowledge of one's own beliefs. On the contrary, if the

theory is correct, knowledge of some of one's own beliefs explains how one can have knowledge of one's own actions. It should be pointed out that this does not mean that the above argument is incorrect. It just means that *not all* our knowledge of our own beliefs can be a consequence of the knowledge we have of our actions of avowing and endorsing them. There must be at least some beliefs that are known independently and prior to our knowledge of our own actions. But if such knowledge of beliefs is possible and, indeed, required by knowledge of our own actions, then it is, of course, not very plausible to think that first person authority in knowledge of one's own beliefs should be explained by it. (Consider also: clearly, that knowledge of belief which is presupposed by the theory must be at least as authoritative as the knowledge that is supposed to result from it. And this prior authority has not been explained at all.)

Is this a good objection to the argument as I have reconstructed it above? The answer depends on how we understand the steps from (3) to (4) and from (8) to (9). In what sense does the primary reason of an action "include" a description of the action? And why should it follow from this fact that a consideration of the primary reason produces (or involves) knowledge of the action? Unfortunately, neither Moran nor Moya are precise on this point. Moran speaks of a description of the action which is known in knowing the primary reason of the action. The description presents the action as choiceworthy by the agent. So knowing the primary reason of my action means (or at least involves) knowing that the action is choiceworthy (for me). But how do I get from this fact (that the action is choiceworthy) to knowledge that I am performing this action? Moran seems to be silent on this point. I will come back to this problem at the end of my paper. For now, we should note that Moran does *not* seem to require the agent to *self-ascribe* a belief in order to know her action. What is important in "knowing the primary reason" of her action, according to Moran, is that the agent knows that her action is choiceworthy or "aiming at some good to be achieved" (Moran 2001, p. 126). Knowing that an action is choiceworthy is not the same, and does not necessarily seem to involve, knowing that *I believe* that such-and-such (e.g. knowing that I believe that the action is choiceworthy).

The same holds if we apply, as Moya suggests, Moran's ideas to the "action" of avowing and endorsing a belief. Such an action, if intentional, must have a primary reason which presents the action as choiceworthy or aiming at some good to be achieved. So in knowing my primary reason for avowing and endorsing a belief I know that avowing and endorsing it is choiceworthy. What makes this kind of

action choiceworthy might be the fact that the belief to be avowed seems to be true; and knowing the action to be choiceworthy might involve knowing some such fact. Again, it is not entirely clear how we can get from this knowledge (that the action of avowing and endorsing some belief is choiceworthy) to knowledge of the fact that we, actually, *have* this belief (the step from (8) to (9)). It should also be noted that this step is not perfectly analogous with the earlier step from knowing that some action is choiceworthy to knowing that I am performing the action. For here (step (9)) it is not to be concluded that I am avowing and endorsing some belief, i.e. that I am performing the action in question. Rather, the conclusion is supposed to be that I have this belief. So the “parallelism between action and belief” (Moya 2006, p. 12) would not be perfect in this reconstruction of the argument. The more important point here is that, again, the argument does not seem to require the agent to *self-ascribe* a belief as a precondition for the conclusion. Knowing the primary reason for avowing and endorsing a belief involves knowing that such avowing and endorsing is choiceworthy, where this might mean knowing that the belief to be avowed is true. There is no self-ascription of belief in such knowledge. It is not the case that, in knowing that the avowal of some belief would be an avowal of a true belief, I know that *I believe* that such-and-such.

If this interpretation of the argument that Moya ascribes to Moran is correct, then Moya’s circularity objection fails. When Moran says that an agent knows the primary reason of her action, he does not mean that she knows that she has some particular belief (and, perhaps, a particular desire). Rather, he means that she knows that the action has a particular quality, namely that of being choiceworthy. There is no circularity in the assumption that such knowledge is a precondition for knowledge of one’s own beliefs.

Moya considers some such response to his objection. His reply is the following: Suppose we take it that by “knowing one’s primary reason” Moran means knowing about some state of affairs (such as that a particular action is choiceworthy). Now, even if we do not have to know that *we believe* that the state of affairs obtains Moya observes that we still have to know the *content* of our belief, namely that the state of affairs obtains. Otherwise the belief could not figure in our deliberation. (Moya’s reason for holding this seems to be that beliefs can figure as reasons in our deliberation even if they are false.) But knowing the content of a belief is also a way of knowing one’s own belief, hence a piece of self-knowledge. Moran’s argument therefore remains circular, since it presupposes knowledge of one’s

beliefs (in the guise of knowing their contents) in order to explain such knowledge:

[I]t seems that our knowledge of those beliefs we form through deliberation is no less immediate and authoritative if we proceed on the basis of facts or truths than if we rely on falsities, provided that we sincerely take them to be truths. If we *do not* take them to be truths, or probable truths, we just do not use them as reasons. But taking something to be (probably) true is just to believe it. States of affairs, then, function in our deliberation as objects or contents of beliefs. So, even if self-knowledge concerning a belief, *B*, which results from deliberation, did not require us to know that we *believe* that *P* as a reason for *B*, but only to know (or believe we know) *that P*, self-knowledge would still be assumed concerning *what we believe*, that is, concerning the *content* of our beliefs. Now, since beliefs have their contents essentially, this amount of self-knowledge would be enough for the circularity objection to remain firmly in place. (Moya 2006, p. 14)

It seems to me that this reply is hardly convincing. First, whether or not the agent's reasons are true does not seem relevant to what is at issue. Her reason for choosing an action is: "The action is choiceworthy". If it turns out later that the action is not choiceworthy this does not mean that her actual reason was: "*I believe* the action is choiceworthy". It just means that she deliberated on the basis of a state of affairs that did not obtain. Saying that she knows her reason is saying that she sincerely believes, "The action is choiceworthy" (even if, in fact, it is not). Perhaps it also means having a belief such as "I will φ *because* φ ing is choiceworthy". It does not mean that the agent believes, "I believe that the action is choiceworthy" or "I take it that the action is choiceworthy". No such second-order belief seems necessary for it to be the case that she knows her reason in Moran's sense.

Moya says that "states of affairs function in our deliberation as objects or contents of belief". And he seems to infer that deliberating on the basis of knowledge (or supposed knowledge) that *P* presupposes self-knowledge concerning the content of our belief that *P*. But is it impossible to deliberate on the basis of (supposed) knowledge that *P* without knowing that one believes that *P*? Why should this be? It seems that for such deliberation to take place it should be sufficient that in my decisions about the matter at hand I am moved

by my (supposed) knowledge. No second-order knowledge seems necessary for such deliberation.⁴

In general, it seems that whenever I know that P I thereby know a content of a belief of mine, namely that P. For how could I know that P without knowing the content that P? But is this *self*-knowledge? Rather, the claim seems to amount to no more than a tautology: By knowing that it is raining I know a content of a belief of mine, namely that it is raining. Saying that knowledge of the contents of one's beliefs amounts to self-knowledge conflates the distinction between knowing that P and knowing that one believes that P. It seems that more argument would be needed to establish such a claim.

On the other hand, we might accept Moya's reply and still find Moran's argument, as reconstructed above, interesting. For the kind of "self-knowledge" that, according to Moya, is involved in knowing the contents of one's beliefs seems to be much less substantial than that expressed in a full self-ascription of belief. If Moran's argument established that such self-ascriptions are possible on the basis of knowing that an action is choiceworthy (hence, according to Moya, on the basis of knowing the content of one's belief that the action is choiceworthy) this would still be an interesting result. It would show how we can come to have interesting self-knowledge (full self-ascriptions of belief) on the basis of a much less substantial and more ubiquitous kind of "self-knowledge" (mere knowledge of the contents of one's beliefs).⁵

It seems to me that the above argument fails for reasons other than circularity. One is that the steps from (3) to (4) and from (8) to (9) are not elucidated. Another reason is that propositions (1) and (6) do not seem to be true for many intentional actions and avowals of belief. Not all our intentional actions are based on processes of deliberation and neither are all our beliefs and avowals of belief. Moya says, rightly, that "a great deal of our beliefs is not formed through deliberation" (Moya 2006, p. 14). For example, we can come

⁴ Sydney Shoemaker has argued that rational deliberation of the kind we engage in to adjust our beliefs to new information requires self-knowledge (for example in Shoemaker 1988 and 1990; cf. also McGinn 1982, p. 20, and Burge 1996). However, his arguments are controversial (cf. Moran's own comments in Moran 2001, p. 112, also Fricke 2012) and Moya does not allude to them.

⁵ At the beginning of his article, Moya describes the "old problem of self-knowledge", which Moran aims to tackle, as the problem of accounting for the authoritative and immediate character of "judgements about one's own beliefs, intentions and other mental states" (Moya 2006, p. 3). Clearly, this notion of self-knowledge is more substantial than that involved in mere "self-knowledge [...]" concerning the *content* of our beliefs" (p. 14).

to have beliefs simply as a result of perceiving something, without reasoning about it beforehand. Yet, such beliefs seem to be known by us as authoritatively and immediately as those beliefs formed through explicit deliberation. If this is correct, then the argument as reconstructed above does not provide an explanation of such authority. As an explanation of first person authority the argument is incomplete because it leaves out a great number, and perhaps the majority, of cases to be explained (cf. Moya 2006, pp. 14 ff.).

However, not only is the above reconstruction of Moran's argument of dubious strength. It is also questionable as an interpretation of Moran's texts. In Moran 2003 we read:

[T]he question about a particular person's future behavior is treated as transparent to a question about what's attractive, or reasonable, or fastest, or otherwise decisively good enough about one of the options. [...] As with Transparency with respect to questions about one's belief, it seems this can only be legitimate if the agent is entitled to assume that the question of what she will do is in fact determined by her assessment of the reasons in favor of a course of action. (Moran 2003, pp. 411 f.)

Here, Moran tells us that the question of what legitimises transparency where knowledge of actions is concerned is exactly analogous to what legitimises transparency where knowledge of beliefs is concerned. I can know what I will do on the basis of a consideration of what is attractive, reasonable, or otherwise decisively good, i.e. "choiceworthy" in the terminology of the earlier quotation from Moran 2001 (p. 126). The question of what I will do (a question about myself) is transparent to the question of what is attractive, reasonable or otherwise decisively good (a question about the world). This transparency is legitimate, Moran says, because of an assumption, which again is analogous to the earlier Transcendental Assumption of Rational Thought: I must assume that what I will do is determined by my deliberation about what is the best course of action.

This quote suggests that knowledge of one's own actions is not different from knowledge of one's own beliefs. It is, rather, another example of transparency that has an analogous foundation. Within Moran's theory, then, the case of knowledge of one's own actions seems to have primarily illustrative functions. Perhaps Moran thinks that in the case of actions it is easier to see that there is transparency and what legitimises it. But he does not seem to regard this case

as more fundamental or as providing further details of the explanation he has in mind for both knowledge of one's own actions and knowledge of one's own beliefs.⁶

But why does Carlos Moya think that "some additional explanation" (Moya 2006, p. 9) is needed of our capacity for authoritatively self-attributing beliefs? His reasoning seems to be that deliberation does not, of itself, necessarily produce self-knowledge. It does determine our beliefs, as the Transcendental Assumption says; but often it does not result in self-knowledge. So what explains self-knowledge, in addition to deliberation about the world? This is the question that, according to Moya, Moran tries to answer with his reference to knowledge of our own actions. We have seen that this interpretation is not very plausible. We should also take into account the following assertion of Moran's: "Suppose you're trying to answer the question whether P, and then deliberate to the answer that yes, indeed, P; nothing in my account explains how we get from this knowledge that P to the self-knowledge that I believe that P" (Moran 2003, p. 409). In this passage, Moran explicitly says that his account does not explain how deliberation about the world produces self-knowledge. It does not explain how we get from knowledge that P to knowledge that we believe that P. It seems that Moya is trying to find a way of filling this explanatory gap with the resources provided by Moran's book.

However, if Moran himself acknowledges that his account is not about how we get from knowledge that P to knowledge that we believe that P, then there does not seem to be much point in trying to find a theory about this in his writings. Why does Moran not try to explain how we make this step? It seems to me that the reason lies in the distinction I made earlier between epistemological and metaphysical explanations of self-knowledge (cf. Shoemaker 2003, p. 400). Moran's account is an epistemological one. He asks what *legitimises* transparency. In such an account it is not necessary to explain how we get from knowledge that P to knowledge that we believe that P. It is sufficient to explain why we have the epistemic *right* to get there. And the epistemic right, when we have it, lies in the truth of Moran's Transcendental Assumption of Rational Thought. Our deliberation determines our beliefs. Therefore, when, in deliberation,

⁶ Of course, Moran is known for his emphasis of notions such as "epistemic agency" and "commitment" in his account of self-knowledge. In my reconstruction, these notions are part of the idea that our deliberation determines our beliefs and actions. My suggestion is that, in Moran's theory, this "Transcendental Assumption" rules our self-ascriptions of beliefs just as much as that of our actions.

we come to the conclusion that P, we have the epistemic right also to assert that we believe that P.

The further question of how we get from the belief that P to the belief that we believe that P is a question about the *metaphysics* of self-knowledge.⁷ It has to do with the mechanisms that underlie the self-attribution of belief. It seems to me that Moya is right to ask for some such explanation. The question of how transparency is possible has this metaphysical dimension to it. But it also seems clear to me that if the self-attribution of belief requires some such metaphysical explanation, then so does the knowledge of our own actions. Not only does Moran not explain how we get from knowledge that P to knowledge that we believe that P. He also does not explain how we get from knowledge that φ is the most attractive, reasonable or otherwise decisively good action to *self*-knowledge that we are φ -ing or will φ .⁸

REFERENCES

- Anscombe, G.E.M., 1963 [1957], *Intention*, 2nd. ed., Blackwell, Oxford.
- Boyle, M., 2011, “Transparent Self-Knowledge”, *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 85, no 1, pp. 223–241.
- Burge, T., 1996, “Our Entitlement to Self-Knowledge”, *Proceedings of the Aristotelian Society*, vol. 96, pp. 91–116.
- Byrne, A., 2005, “Introspection”, *Philosophical Topics*, vol. 31, no. 1, pp. 79–104.
- Evans, G., 1982, *The Varieties of Reference*, Clarendon Press, Oxford.
- Fricke, M.F., 2012, “Racionalidad y autoconocimiento en Shoemaker”, in P. Stepanenko (ed.), *La primera persona y sus percepciones*, CEPHCIS-UNAM, Mérida, pp. 53–73.

⁷Some epistemological accounts of the authority of self-ascriptions of belief might regard the epistemology as incomplete as long as it does not include an account of how such self-knowledge can come about. On such a view, the epistemological account comprises the metaphysical one and the distinction between the two collapses (cf. note 3 above).

⁸The main ideas of this text were first presented at the First Latin-American Analytic Philosophy Conference ALFAN ($\alpha 1$), held in Mérida (Yucatán) in 2010, and I am grateful for the comments I received on that occasion, which, I believe, came exclusively from Olbeth Hansberg. I would also like to thank two anonymous referees for *Crítica* for very useful comments on an earlier version of this text. During my work I received financial support from the Consejo Nacional de Ciencia y Tecnología in Mexico (project “La autoridad de la primera persona”, CONACYT 90167) and from the Programa de Apoyo a Proyectos de Investigación en Innovación Tecnológica of the National Autonomous University of Mexico (project “La primera persona y sus percepciones”, PAPIIT IN400508). I am grateful to both.

- Heal, J., 2004, "Moran's *Authority and Estrangement*", *Philosophy and Phenomenological Research*, vol. 69, no. 2, pp. 427–432.
- McGinn, C., 1982, *The Character of Mind*, Oxford University Press, Oxford.
- Moran, R., 2003, "Responses to O'Brien and Shoemaker", *European Journal of Philosophy*, vol. 11, no. 3, pp. 401–419.
- , 2001, *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton University Press, Princeton.
- Moya, C., 2006, "Moran on Self-Knowledge, Agency and Responsibility", *Crítica*, vol. 38, no. 114, pp. 3–20.
- O'Brien, L., 2003, "Moran on Agency and Self-Knowledge", *European Journal of Philosophy*, vol. 11, pp. 375–390.
- Shoemaker, S., 2003, "Moran on Self-Knowledge", *European Journal of Philosophy*, vol. 11, no. 3, pp. 391–401.
- , 1990, "First-Person Access", *Philosophical Perspectives*, vol. 4, pp. 187–214.
- , 1988, "On Knowing One's Own Mind", *Philosophical Perspectives*, vol. 2, pp. 183–209.

Received: October 15, 2010; revised: May 10, 2012; accepted: July 26, 2012.