

Final Version: O'Brien, L. (2003), *European Journal of Philosophy*, 11: 375-390.

***Moran on Agency and Self-knowledge*¹**

Authority and Estrangement: An Essay on Self-Knowledge, by Richard Moran
(Princeton NJ: Princeton University Press 2001) pp. 201; ISBN: 0-691-08944-2
(Cloth) 0-691-08945-0 (Paper).

1. Knowledge of our beliefs

Self-knowledge is not just another epistemic acquisition, like knowledge of trains or stamps. It is knowledge that lies at the core of our understanding of what it is to be, and what is important in being, a person. Just as fascinating as the centrality of self-knowledge is its fragility. Perhaps the central most notable achievement of Moran's book is the substance it gives to these truths. It provides a rich, illuminating and nuanced discussion of the character and role of self-knowledge.

Moran's central conviction is that the special nature of self-knowledge depends upon it being a product of rational agency. He juxtaposes this view to that of self-knowledge as a 'purely epistemic' gain attained via reasons or evidence. Persons, on the view that emerges, are beings capable of taking both theoretical and deliberative attitudes toward themselves. In general, we will expect the explanation of how we know something to draw on our capacities for theoretical enquiry, for gathering evidence or reasons in favour of that thing. It is therefore natural to try to explain self-knowledge as a product of a particular kind of theoretical enquiry. However, such attempts have proved highly unsuccessful. Moran follows others in finding any attempt to explain the gain of ordinary self-knowledge in terms either of some internal perception, or evidence gained from our observation of our own behaviour as wanting. Such attempts fail to respect the nature of self-knowledge as normally

immediate and authoritative. The failure to provide a satisfactory account of self-knowledge as the result of normal theoretical enquiry has, naturally enough, led philosophers to believe that such attempts are the result of chasing shadows. It has become prevalent to hold that there is in fact no genuine epistemic achievement that is indicated by what is called '*self-knowledge*', and which is expressed in self-ascriptions of the form 'I believe that *P*' (or 'I φ that *P*'). What is going on is explained in a number of different ways. Certain constitutivists explain such seemingly groundless self-knowledge as an artefact of its being part of the concept of belief that we hold a subject who self-ascribes their belief to have the belief self-ascribed. Self-ascriptions are supposed in some way to bring about the facts that ensure their own satisfaction. Expressivists hold that first person self-ascriptions are not really ascriptions at all, but rather are *only* expressions of the first-order belief that *P*. They therefore deny that such self-ascriptions express any epistemological achievement that stands in need of explanation. Moran's criticisms of these views are entirely welcome and successful. In them he brings out his commitment to what I take to be *desiderata*, if indeed not constraints, on an adequate account of how a subject knows her own mental states and acts. These are that the account construe the known mental states and acts realistically and as independent of their self-ascription, and that the account deliver genuine epistemic standing to the knower.

Moran's discussion very effectively explores the connections between immediate self-knowledge and rational agency. His main strategy in seeking to establish the relation between the two is to show what happens to our understanding of the subject as a rational agent, capable of making her mind up on the basis of deliberation, if we suppose her to have knowledge of her mental states only through evidence. In the case

of belief, for example, the business of gathering evidence that one is in a given belief state will result in a de-coupling of the subject's commitment to the truth of the fact that she believes that *P* from her commitment to the truth of *P*. Her evidence via an inward glance, or via observation of her own behaviour, may give her evidence in favour of the truth of 'I believe that *P*' but it does not give her evidence to the same degree for the truth of *P*. The disjoint nature of the sets of reasons in favour of 'I believe that *P*' and '*P*' respectively leave it an open possibility *for the subject* that her judgement as to whether *P*, and her judgement as to whether she believes that *P*, could diverge. However, it cannot in general be an open possibility for the subject that these diverge, since her judgement as to whether *P* is by conceptual necessity *determinative* of her belief with respect to *P*. To understand beliefs as rationally held states which aim at the truth, is to understand them as states formed and maintained by an agent on the basis of her reasons for *P*. This means that the self-ascription of her belief that *P* must in general meet what Moran calls 'the transparency condition' for belief: A self-ascription meets the transparency condition when it is made by consideration of the facts about *P* itself. Meeting the transparency condition is something that no self-ascription made on the basis of evidence can do. A self-ascription that meets the transparency condition is seen to be a *knowledgeable* one on the grounds that a subject's actively making her mind up as to whether *P* is sufficient to give her knowledge that she believes that *P*, if she does.

Moran's work in undermining the ability of standard epistemic 'theoretical' models to capture the nature, role and importance of self-knowledge is convincing and illuminating. I also share the view that the particular character of our knowledge of our beliefs and judgements (as well as our intentions and actions) is a product of our

rational agency, and Moran has given us ample reason to believe that the exercise of rational agency does, and must, give us knowledge without any need for theoretical investigations. What is less clear to me is what Moran's account is of precisely *how* agency gives us knowledge, of what *epistemic* model is being employed to replace the ones that proved inadequate.

It not obvious that Moran is even aiming to give any such account. There are suggestions that Moran takes his discussion to have led us away from epistemic concerns. The book plays throughout on a distinction between theoretical and deliberative stances. From a theoretical stance I ask the question 'what do I believe?' and from the deliberative stance I ask the question 'what to believe?' We are urged to consider self-knowledge as something that comes from the deliberative stance. We are also given a distinction between authority understood as 'knowing best' and authority understand as 'being up to me' and told that:

a person is credited with first person authority when we take the question of what he does believe to be settled by his decision as to what he is to believe. (p. 134)

It can seem at points that the theorist about self-knowledge is being urged to stop wondering about how it is that the person who self-ascribes her own belief is in a particularly secure epistemic position, and rather note that her self-ascriptions have authority in the sense that they are up to her. But it cannot be right to do that. If we are to claim that the self-knowledge attained via rational agency really is *knowledge*, then it is incumbent upon us to make it clear in virtue of what it is being so considered. We need to know why a person who decides as to what he is to believe has authority not only in the sense that his beliefs are up to him, but also in the sense that he knows

what they are. Of course, if there is no such account forthcoming the suspicion will emerge that while the account is not deflationist with respect to the important nature of the phenomenon of self-knowledge, it is deflationist with respect to the claim that it is knowledge. Given Moran's commitment to substantialism we must see him as attempting to offer us not just a characterisation of the phenomenon, its role and its importance but also an explanation of what makes it *knowledge*.

In Moran's explanation of why we can still take avowals of 'I believe that *P*' made on the basis of considering our reasons for *P*, to be reports about ourselves he says:

Any understanding of belief that provides for the minimal idea that believing involves "holding true" will entail that it is at least possible to announce one's belief by reporting on the truth as one sees it. If my intention is to report on my belief as such, and I know (how could I fail to know?) that my belief about *X* is what I hold to be true of *X*, then my intention will not be thwarted if I make this report by considering what is true of *X*. (p. 105)

Assuming that our reports are taken to be expressive of our knowledge of our beliefs, we might take this to be a suggestion as to how a subject comes to have knowledge of her beliefs along the following lines:

- (i) A subject *S* knows in virtue of understanding her concepts that 'a subject's belief about *P* is what they hold to be true with respect to *P*.
- (ii) *S* knows by considering what is true of *P*, that she holds *P* to be true.
- (iii) So, *S* knows that she believes that *P*.

This may be an excellent way for a subject who has the requisite conceptual sophistication to work out what her beliefs are. However, it is clear that such an account will not obviate the need for, but will rather depend upon, an explanation of

why (ii) is true: of *how* S knows that she holds *P* to be true by considering what is true of *P*.

It is clear that Moran's answer to this question must lie in his view of the relation between belief and deliberation. Beliefs are not merely states that represent the world as being a given way, they are states that a subject *herself* forms and maintains relative to her reasons for taking the world to be that way. Furthermore, a subject who reaches beliefs via deliberation is supposed by Moran *to take them* to be settled by her deliberation.

one must see one's deliberation as the expression and development of one's belief and will, not an activity one pursues in the hope that it will have some influence on one's eventual belief and will. ... Were it generally the case that the conclusion of his deliberation about what to think about something left it still open for him that he does in fact think about it, it would be quite unclear what he takes himself to be doing in deliberating. It would be unclear what reason was left to call it deliberating if its conclusion did not count as making up his mind. (p. 94)

Moran's account of self-knowledge seems to draw on two things: the fact that beliefs are reached via deliberation and the fact that the deliberating subject takes them to be so settled. A subject, according to Moran, can answer the question 'what do I believe?' because she can answer the question 'what shall I believe?' knowing that she has the capacity to alter her states in accordance with her reasons. We might take her as having something like the following in mind:

- (i) S knows, by knowing her concepts, that deliberation is such that concluding that there is good reason to hold *P* to be true normally results in one holding *P* to be true.
- (ii) As a result of deliberation S concludes that she has good reason to hold *P* to be true.

- (iii) S knows that she concludes that there is good reason to hold *P* to be true.
- (iv) She knows that she holds *P* to be true.
- (iv) A subject S knows in virtue of understanding her concepts that a subject's belief about *P* is what they hold to be true with respect to *P*.
- (v) S knows that she believes that *P*.

On this construal of self-knowledge, our knowledge of what we actually hold to be true depends upon our knowledge of what we conclude there is reason to hold true. However there are problems with giving an account of what we actually hold to be true via an account of what we take ourselves to have good reason to hold to be true. This account will deliver knowledge when a subject judges in accordance with what she considers she has reason to judge. But, the account would seem to leave it an open possibility that the subject can judge without knowing it (when she judges without a prior conclusion to the effect that she has good reason to judge it). Also, it leaves it open that a subject can have warrant for thinking that she has judged that *P* when she has in fact only concluded that she has good reason to judge that *P* and then drawn back in fit of epistemic risk aversion. Further, the account still owes us a story about how a subject knows her conclusions about what she has reason to believe and one is inclined to think that any satisfactory explanation for that can be applied directly to the case of belief.

However, it would not be right to take Moran's account to be represented by the above.² The immediacy of self-knowledge, that he takes to be illuminated by considering beliefs as the products of deliberation, suggests that we should not see deliberation as a process whereby a subject decides what she should believe and then

acts on her decision. Furthermore, the immediacy of self-knowledge is in tension with the idea that such knowledge is mediated by the subject's belief in a generalisation to the effect that 'the conclusions a subject reaches about what to hold true generally result in the subject holding it true'. The immediacy and authority of self-knowledge is, for Moran, due to the fact that a belief itself (rather than some reason to believe) is the conclusion of deliberation, and understood to be so by the believing subject. Deliberation, for Moran, takes place against a supposition by the subject that by deliberating *she can immediately* make up her mind and in a way consonant with the pattern of her deliberation.

I think that Moran's idea is rather that, given the subject's understanding of deliberation, the subject knows that all that needs to be done to find out what she believes is that she reach a conclusion, via deliberation, about what is the case. The conclusion reached by deliberation will be her belief, and she knows this, given her understanding of what deliberation is. The account should be thought to involve a *practical procedure* that a subject who knows her concepts, and the connections between them, can apply in order to come to have knowledge of her beliefs. In understanding the direct connections between deliberation as to whether *P* and resulting beliefs with respect to the true of *P*, the subject can come to know what she believes by answering the question whether *P* and self-ascribing her answer. By knowing her concepts, she knows that she can find out whether she believes that *P* is true by calculating whether *P* is true and adding 'I believe that...' to '*P*' if it is.³

If this is the account then it I think it has much going for it. In particular I think it is right to take the subject's actually coming to hold *P* to be true on the basis of reason to

be part of the warrant for the knowledge. However, I have some concerns. First a question: while I think Moran is right to think that we need an account of our knowledge of our perceptions that is different from our account of our beliefs, we might expect an account of our knowledge of our beliefs to explain our knowledge of other non-experiential cognitive attitudes such as our own assumptions and suppositions. An account of how to treat less committed cognitive attitudes would be welcome. Second a worry: it seems that too great a role is being played by the subject's reflections on the nature of deliberation. It is an important fact that the capacity for deliberation, and for holding and revising beliefs on the basis of reasons, does not seem to require a subject to have full grasp of the concept of belief. Consider Peacocke's example of a small child who has no full conceptual mastery of the concepts 'belief', 'deliberation' etc. but who wonders where his toy is.⁴ He considers that it is most likely that the toy is in the cupboard on the grounds that that is where it is usually put away. He then remembers that Auntie is staying and keeps putting things in the wrong places, and so revises his belief that the toy is in the cupboard. He instead reaches the conclusion that it could be anywhere. The child can clearly be seen to be deliberating over the question 'where is my toy?', and as such is deliberating about what to believe. But he cannot be thought to be doing so under that intention or description. There is, however, the strong intuition that the child has all that is needed by way of grounds in order to know what his beliefs are. Consider a stage when the child has a partial grasp of the concept of belief. He is not good at passing the false belief test, and has not the conceptual sophistication to form beliefs about his capacity for deliberation nor about the connection between his wondering where his toy is and his belief about where his toy is. Nevertheless, he happily expresses his conclusion by saying 'I believe that teddy could be anywhere'. Would we not want to say that he

nonetheless has immediate and authoritative knowledge of his own mind? It seems to me that we would want to say that he is aware of his beliefs in just the way we are, but that he has not the concepts to articulate his grounds for that awareness. If that is so then it is not clear that we want to take the subject's explicit knowledge of the nature of deliberation and belief to function as part of the warrant for self-knowledge.

The difficulty in determining *how* Moran thinks rational agency gives us knowledge of its product comes to a significant extent from there being little guidance as to what positive conditions he places on knowledge. We know that Moran does not take traditional forms of justification by observation as required for knowledge, and it is clear that he thinks mere reliability is insufficient for knowledge, at least in this case. However, there is less guidance as to what *is* required. If we are right about the account we have attributed to him, then while a subject's actually deliberating and forming a belief as a result will play some epistemic role, it only does so via her knowledge of what deliberation involves. Both her forming a belief and her reflective understanding of the process of forming her belief seems to be required as a source for warrant.

My own view is that the subject herself need have no reflective understanding of her concepts or cognitive capacities in order to be credited with knowledge of her beliefs, but I agree with Moran that seeing beliefs as the exercise of rational agency is critical to our having immediate and authoritative self-knowledge with respect to them. In my view we can see the subject's knowledge as resting on an entitlement (of the kind articulated by Burge, Dretske, Peacocke) which she herself may be quite unable to articulate or understand.⁵ Of course, what account we give of the source of that

entitlement will differ. We may take it that the conceptual truths Moran brings out with respect to the nature of deliberation and its connection with belief formation, maintenance and revision, could by themselves be a source for the subject's entitlement to self-knowledge. The suggestion would be as follows:

- (i) The concepts involved in, and nature of, deliberation are such that what S holds to be true with respect to P is determined immediately and consonant with S's consideration of what is true with respect to P .
- (ii) S considers P to be true
- (iii) S has knowledge, via entitlement, that she holds P to be true.
- (vi) The concepts of belief and holding true are such that a subject's belief about P is what they hold to be true of P .
- (iv) S has knowledge, via entitlement, that she believes P to be true.

The problem faced by such an account of the source of the entitlement, in common with all purely externalistic accounts of warrant, is that it needs to make it plausible that a subject can be granted knowledge that *she ϕ -ed that P* , simply by ϕ -ing that P . How does she have knowledge given that P is the object of her attitude, rather than herself, and she has no access to the conceptual truths connecting her ϕ -ing that P and P ?⁶

I would favour trying to work out some middle position by articulating a kind of practical knowledge or awareness a subject has as a deliberator which acts as source for the subject's entitlement. Although Moran does not mention the possibility of such

practical knowledge, it might be that he too would be happy with this prospect. On such a story something like the following is true:

- (i) A subject's coming to have knowledge of her own beliefs via deliberation rests on an entitlement to self-ascribe her belief.
- (ii) The entitlement rests on some knowledge on the part of subject that she is deliberating between options for *her* led by reasons.
- (iii) Such knowledge consists not in the subject's her own conceptual mastery of the pre-suppositions of deliberation, but rather in a non-conceptual practical awareness manifest in the subject's capacity to be led by reasons to judgement.

The idea would be that a pre-supposition of deliberation is a practical knowledge of the options available to me. Of course making out such an account depends on making out the thought that there can be practical knowledge and awareness which outstrips the subject's capacity to conceptualise it. Such an account would steer a course between the conceptualist, as we have called him, who thinks that a subject's warrant for self-knowledge depends to a significant extent on her own mastery of the conceptual connections, and the austere entitlement theorist, let us call him, who think that the warrant is a purely external one with the conceptual connections doing all the work.⁷

Knowledge of our Actions

Moran holds that knowledge of our actions and our beliefs are in important respects the same. I think he is right. However, the task of reaching a sharp understanding of what model of knowledge via agency Moran is giving is in fact made a little harder by Moran's account of how it is we know our own physical actions. The explanation he *prima facie* gives of our knowledge of our actions, via knowledge of our intentions,

has problems and does not carry over well to the belief case. His suggestion is, and here he is following Anscombe, that we have non-observational knowledge of our actions by knowing our primary intentions in carrying them out. An agent can have authority over her actions because her actions are based on intentions which are based on reasons for doing one thing rather than another. As Moran puts it:

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. (p. 126)

For an agent to conceive of himself as capable of forming an intention and implementing it (which, I take it, is necessary to conceiving of himself as an agent at all) he must take his intentional action to be determined by his reasons and thus he is in a position to know a true description of his action in knowing his reasons. (p. 127)

A number of worries come to mind. First, it is not clear how helpful knowledge of our intentions is likely to be in explaining our knowledge of our actions. We often, all too often as Moran fully realises and explores, form intentions to act but fail to act on them. I intend to go to sleep directly after I have finished the current chapter of my book. I form the intention on the basis of reasons sincerely held, I have to be up at 7.00am, I have to finish a critical review, I cannot think well when I have had too little sleep and so on. I get to the end of the chapter and just cannot bring myself to put the book down. I know immediately and without observation that I am reading on even without having intended to. And had I managed to stop reading I might have found myself pleased, and maybe a little surprised, that I had acted on my intention. These considerations suggest that knowing my intentions is not going to amount to a very secure way of coming to know my actions. Unless I have independent assurance that I am in a particularly determined state of mind the ever present possibility that

this intention will go un-acted on must surely undermine my authority over my actions.⁸

Second, we all too often act on our intentions and fail to satisfy them. Consider, how often we intend to do things we are not at all confident of pulling off. I strive to pot the pink ball on a full-size snooker table. My success rate is utterly miserable and I know it to be – I know it to be a long shot – but this does not stop me intending and trying to pot the pink. However, in order to know that I have in fact potted the pink I have to peer down to the other end of the table to see whether it went in the pocket.

The gap between our intention and our actions, which these cases bring out, means that our authority over our intentions, which can be given the same explanation as the authority of our beliefs, does not help us to explain the authority we have over our actions. The story for beliefs rested on the fact that beliefs are spontaneous and immediate changes of mind brought about by deliberation. Intentions may also be direct products of deliberation in this way and so knowable in the way that beliefs are. But if we see actions as brought about by prior intentions, and see our knowledge of actions as accruing from our knowledge of our intentions, then our knowledge of our actions cannot be as secure as it seems to be. One way to remedy this is to see actions themselves, rather than intentions as the standard conclusions of our deliberation. If they are the direct products of deliberation in the way that beliefs and intentions are they might be knowable in the same way. This is the way I think we should go.⁹

Perhaps Moran will say that he did not mean that we know what we are doing by knowing our *prior intentions* but rather that we know what we are doing by knowing

our *intentions in acting*. Suppose we take prior intentions to be conclusions of practical reasoning stored up for future action. We may hold that we often act without prior intention, but claim that even in those cases where the action itself is the conclusion of a bit of practical reasoning, it will be an action with a primary purpose that reflects the primary reasons for doing it. To the extent that any action is done for reasons held by the agent there will be a description of it that reflects the agents aim in so acting. But if this is so we will still need to know how to deal with the many cases where we act with an intention, perhaps without forming any prior intention, even though we know it unlikely that that action will come off. How can it be that I know without observation that I am potting the pink when I have no confidence that the pink has gone in the pocket? Moran may respond that his claim was not that we have first person authority over our actions *per se*, but only over what we are trying to do, over our intentions in acting. This would be a pity – partly because it is a very much blander claim and partly because it seems false to the phenomenon. It seems to me that our intentional actions *are* such that we are first person authoritative over them, even if not under the description that reflects our primary intention. I may not know whether I have potted the pink but I do surely know immediately and authoritatively that I am leaning over and moving my arms in my effort to pot the pink.¹⁰ But that knowledge does not seem to have been explained by the account via primary intentions.

These considerations bring out the fact that there are important disanalogies between belief and action. In particular, the contents of belief are such that if you cannot think them you cannot intend to think them either. In contrast one can intend to do things

that in fact one cannot do. We need to reflect these differences in our accounts of how we know our beliefs and actions.

Knowledge of our Moral Attitudes

The richness of Moran's book would not be well reflected unless we spend at least some time on his discussion of moral and emotional attitudes. It is a theme of the book, perhaps *the* theme of the book, that although the capacity to be immediately self-conscious with respect to our attitudes is part of what it is to have such an attitude, being self-conscious of our attitudes in other ways, being conscious of them in ways we might be conscious of another's attitudes, can be undermining of the attitude itself. It is part of what it is to be the bearer of an attitude held for reasons that certain kinds of reflection on it can make it disintegrate. The explanation of the potentially undermining nature of such reflection in the case of belief is most clearly explained by the possibility of violating the transparency of belief. I cannot both believe that *P* is true, and take a reflective attitude to my belief that *P* that leaves it open that I doubt that I believe that *P*. If a belief is an essential component of moral attitudes, then such attitudes can be logically undermined by further reflective attitudes that take the first attitudes as their object. A moral attitude is logically undermined if a belief that is part of a first attitude contradicts a belief that is part of the reflective attitude. Suppose that pride involves a belief to the effect that something I have produced is genuinely worthy in a given respect. I cannot both be proud of X in respect Y, and hold that my pride is pitiful given a belief in the worthlessness of X in respect of Y. If I re-assess the worthiness of my product in this way I thereby lose grounds for holding the attitude of pride. I cannot both believe the product to be

worthy and unworthy in the same respect. In such cases there are logical reasons why a subject cannot both have the attitude and have a certain other attitude to it.

But Moran suggests that taking a moral, or emotional, attitude as the object of a second moral, or emotional, attitude can be undermining of the latter, even when there is no contradiction between the beliefs or contents of the commitments involved. For the remainder of this piece I want to explore such undermining.

Moran discusses at some length the following example from Kingsley Amis. Consider a married man walking home after a night of casual infidelity. On his way he describes himself as:

feeling a tremendous rakehell, and not liking myself much for it, and feeling rather a good chap for not liking myself much for it, and not liking myself at all for feeling rather a good chap. (p. 175)

Moran argues that the rakehell is not committed to contradictory beliefs. In feeling himself to be a good chap for feeling shame he is not denying, but is indeed relying on his belief that he did something wrong or cruel. But nevertheless his reflections do seem in some way to undermine his shame. Whilst it is not the central element of his explanation of what is going wrong in the case of the rakehell, there is an aspect of Moran's explanation that I want to press him on. A part of his discussion of how it is that emotions can exclude each other, despite there being no logical conflict between them, appeals to the fact that they are episodic attitudes with an experiential aspect.

Beliefs are not episodic like emotions; they do not typically run their natural course and then fade out like a burst of anger or relief. There is a temporal dimension to the moral meaning of various attitudes. (p. 180)

As such, it seems, they stand to push out or supplant each other. Partly because of this "an emotional attitude constitutes something closer to a total orientation of the self"

(p. 181). Moran amplifies this point with appeal to Hume:

The sense in which a certain moral attitude may involve something like a total perspective on its object is crucial to Hume's account of pride and humility, and relevant to the oscillating quality of the rakehell's reflections. In the treatise Hume says: "'Tis impossible a man can at the same time be both proud and humble; and where he has different reasons for these passions, as frequently happens, the passions either take place alternately; or if they encounter the one annihilates the other, as far as its strength goes, and the remainder only of that, which is superior, continues to operate upon the mind. (n. 18, p. 181)

While a subject who comes to feel some pleasure or comfort in the fact that they feel shame for their wrongdoing is not abandoning their belief that they did wrong, in taking pleasure in their shame they are qualifying their total outlook and pushing aside the feelings of shame. This is undermining of their shame.

But this is misleading. Even if emotions are episodic it is not obvious why this means that they push each other out. Distinct feelings seem to be able to co-exist much as distinct perceptions can. But anyway, emotions do not seem to be episodic in the way suggested. It might be that certain experiential feelings are episodic but the relationship between such feelings and emotions is far from clear. Emotions certainly manifest themselves in feelings, as they do in actions, bodily changes and thoughts but the dimension of the feeling does not even approximately determine the dimension of the emotion.

To take the emotion of shame that is in play in the rakehell example. Shame □ deep, genuine, full-blooded shame □ can last years. A deeply embedded shame can condition almost every action for a subject, the thoughts of that thing of which she is ashamed can give rise to different feelings on different occasions, sometimes to

feelings of sickness and self-loathing, sometimes to restlessness and frantic resolve, the emotion can be manifest in the colour rising on her cheek given certain stimuli, or in her shifting in her chair and so on. There seems to be no single feeling, or action or bodily reaction, or fixed collection of such things, which is the emotion of shame, and the feelings that do manifest the shame can last a longer or shorter time. And the mere oscillation of feeling does not by itself make us doubt the genuineness of the emotion. Had the rakehell oscillated between a feeling of despair that all was lost, and a feeling of optimism that things could be put right, we would not similarly doubt his shame. It is not even clear that the attempt to seek comfort in the *fact* of one's shame need be undermining of the shame. In cases where the shame is truly painful a subject may seek such escape. The mark of its being genuine shame is that the subject will be dragged back to her discomfort. If emotions are dispositions that manifest themselves in a complex variety of ways then even such oscillations of the rakehell could prove consistent with genuine shame. What really makes one doubt the rakehell is not his attempt to escape his dislike of himself, not even his assertion of having for a moment succeeded, but his cheery tone and the suspicion that his escape into such a jolly iteration was complete. Imagine giving the quotation a bit of a 19th century Russian novelist's twist:

feeling myself to be a traitor, and hating myself for it, feeling comforted that at least I hated myself, only to fall back into disgust at the unearned respite.

I think we are much more comfortable with the idea that we have here a candidate for genuine shame.

The view of an emotional attitude as constituting a total orientation of the self which pushes out other emotional attitudes involves a much stronger claim than the narrower point that a subject's taking an outsider's stance to their own moral and emotional

attitudes can be undermining of them. The former, certainly with Hume's gloss on it, suggests a simplicity and well-ordered character to our moral and other emotions which belies the truth. It is not the case that we have one emotion at a time, nor indeed one feeling at a time. Nor is it the case that any seeming mixture of emotions or feelings are the result of one being replaced by another in serial. Rather, having different emotions and feelings at the same time is common and central, indeed it seems to be that it is our normal condition. Hume is perhaps forced by the theory of ideas to hold that we cannot be proud and humble for different reasons at the same time, but surely he is wrong about this. I can and do feel proud of having just learnt to play Frère Jacques on the piano at the same time as feeling humble in the light of all that I might have learnt. The child who feels ashamed of the damage done by some mischief committed may not be able to suppress their simultaneous pride at having been bold enough to commit it. The grief stricken mourner may yet feel comfort in experiencing just how much they loved. When these distinct emotions exist mixed up and side by side they are not undermined, even if they are conditioned by the presence of the others. The possibility of such mixed emotions is pointed out, with her usual ingenuity and insight, by George Eliot. At the end of Chapter 4 of *Daniel Deronda*, speaking of Gwendolen Harleth, she says:¹¹

However, she had the charm, and those who feared her were also fond of her; the fear and the fondness being perhaps both heightened by what may be called the iridescence of her character — the play of various, nay, contrary tendencies. For Macbeth's rhetoric about the impossibility of being many opposite things in the same moment, referred to the clumsy necessities of action and not to the subtler possibilities of feeling. We cannot speak a loyal word and be meanly silent, we cannot kill and not kill in the same moment; but a moment is room wide enough for the loyal and mean desire, for the outlash of murderous thought and the sharp backward stroke of repentance.¹²

So it seems not to be the presence of other emotions and feelings competing for space or attention that in itself undermines our emotional attitudes. Moran's example is,

however, one where one emotion is the object of another. Perhaps the problem is not that we have two or more emotions competing for the total orientation of the self, but rather that we have an emotion as the object of a second emotion. However, I think it is important to note again that the mere presence of the second emotion need not by itself undermine or push out the first emotion. It is not undermining of the primary emotion to be pleased, or relieved at one's delight at seeing someone again, to be anxious at one's excitement, or nervous at one's fear. Nor is there I think a particular problem about having emotions caused by our moral emotions. Given that we can have a view about how we should feel on given occasions, and how we fear our weaknesses might lead us to feel, our actual reactions might cause relief and disappointment in accord with our sense of what is right. I can feel relief at being disgusted rather than amused at the teasing of another, I can feel pleased that I feel sympathy rather than fear at another's hurt. I can feel pleasure and satisfaction at the regard I feel for another.¹³ Judging by Moran's description of our emotional lives in the rest of chapter 5, and indeed throughout the book, there is little in the foregoing remarks on the possibilities of mixed emotion that he will disagree with. However, the view I have been concerned to dispel is suggested by a part of his treatment of the rakehell. It is important to identify it and set it aside from his central explanation.

The explanation of the rakehell that Moran takes to be the central one is that an emotion stands to be undermined, not by a subject taking a stance in relation to their own emotions, but rather by a subject's taking an *outsider's* stance in relation to their own emotions. Taking an outsider's stance can undermine the attitude whether or not her second order attitude has either a purely evaluative element or belief commitments that are in tension with her first order attitude. Whether I view my attitude of

sympathy as evidence for my praiseworthiness, or as an imposition on my plans, or as a sociological construct, or as well expressed, I stand aside from it and so undermine it. It is taking an attitude towards ourselves in the way we might take an attitude towards another that is the problem, not the fact that we have a mixture of emotions, nor that we have emotions as the objects of other emotions.

So, in virtue of what is the rakehell taking an outsider's attitude towards his emotions? The attitudes of shame and pride that seem to be involved, are self-regarding attitudes which take two objects: the subject themselves and the source of their shame or pride. They stand in contrast to other emotional attitudes of sympathy, love, fear that take one object. This obviously has an effect on how they stand to be iterated. There is no bar on self-regarding attitudes surviving the taking of an object, or being taken as the object, of non-self-regarding attitudes: one's pride can make one anxious or one's shame relieved without undermining the pride or shame. One can also be proud of one's emotional or moral attitude to others: one can be ashamed of one's hatred and proud of one's love without thereby undermining the love or hate. However, the rakehell is indulging in the particularly complicated business of taking a self-regarding attitude with a self-regarding attitude as one of its objects. This has the result that, not only is *he* the object of two of his attitudes at once (which need not itself be problematic if the commitments of the attitudes are compatible), but also his evaluation of himself is functioning as grounds for fixing an evaluation of himself. His taking an evaluation of himself as grounds for fixing an evaluation of himself is what constitutes his taking an outsider's view of himself, and is on Moran's view what undermines his shame. In fact, it seems to be a quite general truth that using one's evaluation of X (whether X is oneself or not), as evidence for an evaluation of the

same X, will stand to undermine the status of one's original evaluation, whether the two evaluations are in the same direction or not. In using one's first evaluation as evidence one has dropped one's commitment to it. Suppose I judge a painting beautiful in its use of line and colour. I also judge it ugly in its theme or mood. That seems okay. It is beautiful in one respect but not in another. But now suppose that I judge the painting beautiful in its use of line and colour and then judge the painting ugly in that I, who have terrible taste, have judged it beautiful in line and tone. I have clearly undermined my first evaluation, even though it may be perfectly true that I so judged it, and that I have terrible taste. Now suppose I judge the painting beautiful in line and colour, and then go on to re-affirm my evaluation by judging it beautiful in that it has elicited from *me* a judgement that it is beautiful in line and colour. Again, we come to doubt that the first judgement is genuine. How can my first judgement on the worth of the painting add anything to my view of it unless it was taken by me to have settled the matter, and if it was not sufficient to settle the matter adding it as evidence now will not help. In using the first evaluation as evidence for a further evaluation one removes the force of it *as a standing evaluation*, but its worth as evidence depends upon it being so. What the rakehell does in finding that he likes himself for not liking himself is remove the force of the first evaluation. As Moran puts it:

Reopening the question of how he is to feel undoes the creditworthy aspect of his shame as a settled attitude. The credit for this response depended on the degree to which he was resolved on the question of how he is to feel about himself. Hence, re-opening the question cuts him off from that meaning, and the credit he seeks for it. (p. 186)

This explanation of Moran's is rightly seen as the heart of the problem with the rakehell, and it is an instance of a important point about evaluation in general, the explanation of which lies in the nature of our relation to our own attitudes. It as an

insight that survives, and needs to be set aside from, the temptation to think that problems with co-existing attitudes flow from them having episodic and experiential natures.

Lucy O'Brien
Department of Philosophy
University College London
Gower Street
London WC1E 6BT
United Kingdom

l.o'brien@ucl.ac.uk

NOTES

¹ This paper was first written for an 'author meets critics' session at the APA, Central Division meeting in Cleveland, April 2003. Many thanks to David Velleman for the invitation, to Richard Moran for his response at the APA, and to Mark Sacks for crucial comments and discussion. The paper was written during research leave jointly funded by UCL and by the Arts and Humanities Research Board. I am extremely grateful to both institutions. All page references will be to Moran (2001) unless otherwise specified.

² Moran's account of our knowledge of our actions, and his suggestion that there is a close parallel between our knowledge with respect to belief and action, might, however, lead one to attribute this view to him. We will come to that presently.

³ Consider an analogy. I might know, by knowing my mathematical concepts, that I can find out the length of the hypotenuse of a right-angled triangle by finding the square root of the sum of the squares of the two sides. Given a particular right-angled triangle, and knowing this procedure, I can go through the calculation and express my resulting knowledge of the length of the hypotenuse of this triangle by adding 'the length of the hypotenuse is...' to whatever result I get.

⁴ Peacocke (1996, p. 132)

⁵ See Burge (1993, 1996), Dretske (2000), Peacocke (2002),

⁶ If I calculate the sums of the squares of the two right-angled sides can I know the length of the hypotenuse without knowing Pythagoras' Theorem?

⁷ See O'Brien (forthcoming).

⁸ This point is also made in Owens (2003).

⁹ We can also note that we can draw a parallel distinction between forming an intention and acting on it in the case of belief. The epistemic relations will work out similarly. I might form the intention to believe that *P* if I gain good evidence to the effect that *Q*. As with the action case I think it would be a mistake to explain my knowledge of any resulting belief that *P* on the basis of my knowledge of the intention. There was a strong chance that I would fail to realise my intention, when good evidence for *Q* came along I could easily have found myself in a sceptical mode, thinking that good evidence for *Q* was not sufficient reason to embrace *P*.

¹⁰ For my suggestion about the source of such authority see O'Brien (2003).

¹¹ I did not find the Eliot quotation; it was, very generously, passed on to me by Moran with the suggestion that I might want to use it in this piece.

¹² Eliot (1967, p. 72). The reference to Macbeth is to Act II, Scene 3: "Who can be wise, amazed, temperate, and furious, loyal and neutral, in a moment?"

REFERENCES

Burge, T.(1996), 'Our Entitlement to Self-knowledge', *Proceedings of the Aristotelian Society* 96: 91-116.

(1993) 'Content Preservation', *Philosophical Review* 103: 457-488.

Dretske, F. (2000), 'Entitlement: Epistemic Rights Without Epistemic Duties?' *Philosophy and Phenomenological Research* 60: 591-606.

Eliot, G. (1967) *Daniel Deronda*, Harmondsworth: Penguin.

Moran, R. (2001), *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton NJ: Princeton University Press.

Owens, D. (2003), 'Knowing One's Own Mind' in 'Deliberation and Irrationality in the History of Philosophy', ed. F. Pironet and C. Tappolet, *Dialogue* 42.4.

O'Brien, L. (forthcoming), 'Self-Knowledge, Agency and Force'.

(2003), 'On Knowing One's Own Actions' in Eilan and Roessler (ed.) *Agency and Self-Awareness*, Oxford: OUP.

Peacocke, C.(1996) 'Entitlement, Self-Knowledge and Conceptual Redeployment', *Proceedings of the Aristotelian Society* 96: p. 117-158

(2002), 'Three Principles of Rationalism', *European Journal of Philosophy*, 10: 375-397.