Inquiry, 46, 395–408

First Persons: On Richard Moran's *Authority and Estrangement**

Taylor Carman
Barnard College, Columbia University

Richard Moran's *Authority and Estrangement* offers a subtle and innovative account of self-knowledge that lifts the problem out of the narrow confines of epistemology and into the broader context of practical reasoning and moral psychology. Moran argues convincingly that fundamental self/other asymmetries are essential to our concept of persons. Moreover, the first- and the third-person points of view are systematically interconnected, so that the expression or avowal of one's attitudes constitutes a substantive form of self-knowledge. But while Moran's argument is wide-ranging and compelling, he relies throughout on an overly intellectualized conception of first-person attitudes as attitudes of reflection or deliberation. That conception is at once implausible and unnecessary to the main current of his argument, whose goal is to demonstrate that our self-conception as persons depends on both the distinctness and the interconnectedness of our first- and third-person perspectives on ourselves.

Premodern and early modern philosophers seem to have regarded the soul as a kind of thing, a peculiar kind of thing, the presence of which constituted something's being alive, and which moreover figured as the locus or container of certain peculiar kinds of states or events: perceptions, thoughts, memories, desires, dreams. Descartes at once transformed and continued that tradition, divorcing the soul from all things biological while nevertheless still referring to it as a *res cogitans* in contrast to *res extensa.* The soul was for him not a living or even a physical thing, but a thing nonetheless. And contemporary philosophers of mind persist in worrying about the ontological status of mental states and events, particularly states of consciousness, in what they suppose must be a purely physical world.

The mind-body problem, however, constitutes only one part of the Cartesian legacy. Another, at least equally important part has to do with the peculiar logic of the *cogito*, in which Descartes recognized a radical asymmetry between our immediate first-person access to our own thoughts and our mediated third-person knowledge of bodies and the minds of others. What is strange and interesting about the *cogito* is precisely that it cannot

*Richard Moran, *Authority and Estrangement: An Essay on Self-Knowledge.* Princeton, NJ: Princeton University Press, 2001, xxxviii + 202 pp. $16.95. Unprefixed references are to this book.

Routledge Taylor & Francis Group DOI 10.1080/00201740310002424 © 2003 Taylor & Francis

simply be converted into a *cogitat*: 'I think, therefore I exist' is not a special case of 'X thinks, therefore X exists'. To begin with, any valid inference from 'X thinks' to 'X exists' would require the intervening major premise, 'All thinking things exist', which Descartes explicitly rejects as unnecessary to the insight conceived in the first-person singular.[1] Moreover, no such argument would be sound without some guarantee that 'X thinks' is true. But how am I to know that? *My* thinking '*he* thinks', after all, does not guarantee its truth in the way my thinking '*I* think' does, for 'I think' is *made true* by my thinking it.

Descartes thus almost single-handedly initiated our modern philosophical preoccupation with the self and self-knowledge, and yet recently his influence has been felt almost entirely negatively. The philosophy of the past hundred years or so might fairly be called the Age of Anti-Cartesianism, though in that time there have been nearly as many anti-Cartesianisms as there have been anti-Cartesians. One compelling argument that has been leveled against Descartes is that he erred not simply in drawing an overly sharp distinction between the mental and the physical, though he certainly did that, but rather – indeed precisely *in order* to draw that distinction as sharply as he did – by pressing minds and bodies too much into the same conceptual categories: the categories of thing, property, cause and effect, and so on. It this spirit Gilbert Ryle famously dismissed what he called the 'paramechanical hypothesis' driving substance dualism, the hypothesis that 'Minds are things, but different sorts of things from bodies'.[2]

Kant is the original source of a competing insight to the effect that the kind of perplexity about the self that motivates dualism springs not from the peculiar nature of mental as opposed to physical things, but rather from characteristic conflicts that emerge between two different perspectives we can (and must) adopt with respect to ourselves: the theoretical standpoint and the practical or agent's point of view. Versions of this idea gained momentum in the intervening centuries in thinkers as diverse as Kierkegaard, Wittgenstein, Heidegger, and Sartre, and it finds renewed and innovative expression in Richard Moran's rich and stimulating new book. I will divide my remarks in what follows between what I think is right and important in Moran's project and what I find problematic and unsatisfying.

I

What is right and important, it seems to me, is Moran's guiding idea that traditional problems about the self and self-knowledge can only be understood in light of both the irreducible asymmetry and the complex intermingling of the first-person position we occupy in our engaged practical attitudes and the 'observer's standpoint' or third-person perspective we take

on others, and which we can assume, at least up to a point, with regard to ourselves.

More specifically, one of Moran's most compelling insights is his recognition that the fundamental asymmetry between first- and third-person perspectives runs deeper and wider than any purely epistemological approach to the problem of self-knowledge is able to accommodate. That asymmetry has epistemological implications, to be sure, but it also makes itself felt more broadly in moral psychology, and indeed in everyday life. Moran argues persuasively that both the difference between the two perspectives and their unavoidable interconnectedness are equally essential to our overall self-conception and together constitute our concept of a person. Persons are in principle perspectivally complex, and Moran's book is a profound reminder of how far that complexity extends beyond the exclusively theoretical purview of epistemology.

The epistemologically puzzling fact that Descartes first saw is that my relation to my own experiences and attitudes is fundamentally different from my relation to the experiences and attitudes of others. Of course, Descartes all too abruptly relinquished that insight by conceiving of the mind as an object of theoretical knowledge, literally introspection, a kind of perception directed on an inner domain of mental phenomena radically distinct from external physical nature. Again, conceiving of the mind in that way divides it too sharply from the world precisely by assimilating it too much with other objects of knowledge, and so in effect assimilating all knowledge to a single notion of theoretical contemplation. Still, Descartes's insight persists, in spite of his failure to do it justice. The basis of my knowledge of my own mind is essentially different from the basis of my knowledge of the minds of others, moreover I enjoy an authority about my own thoughts and sensations that is denied to others, whose knowledge of them is necessarily mediated by my testimony and behavior. How are we to understand that first-person form of self-knowledge and its unique epistemic and practical authority?

One way to be anti-Cartesian is simply to deny that there is anything substantive or legitimate about self-knowledge from the first-person point of view. Moran points out, however, that skepticism about substantive first-person self-knowledge often tacitly relies on the 'Perceptual Model' of knowledge that Cartesianism takes for granted. If self-knowledge is supposed to consist in some kind of quasi-perceptual introspection, then it is indeed reasonable to be suspicious of the very idea. But of course the peculiar immediacy of self-knowledge is nothing like the direct perception of an object. Being aware of your own thoughts by *having* them is crucially different from being aware of things distinct from and independent of your awareness of them, for example by *seeing* them. It is a mistake, then, to construe reflexive self-consciousness as a special case of consciousness of *x*, generically conceived. For even complete and perfect theoretical knowledge

of my own beliefs falls short of the kind of self-understanding I have by relating to those beliefs *as my own,* that is, by being in a position to declare or avow them as mine.

What a purely theoretical approach to self-knowledge obscures, Moran argues, is the unique importance it has for the person whose knowledge it is. Knowing one's own mind is not just a matter of having access to some information about oneself, but amounts instead to a kind of well-being that is essential to our concept of a person. More profoundly, in the absence of the kind of self-knowledge I can express or avow, rather than merely attribute to myself from a third-person point of view, it is unclear what would allow me to identify those beliefs as *mine,* as opposed to simply *these.* As Peter Strawson pointed out long ago, the indexical 'these' will be doomed to obscurity unless we understand it to indicate ownership by someone. Without the primitive concept of a person, no demonstrative reference to a mere Humean bundle of ideas would even be possible.[3] The epistemology of the first person thus stands in need of an adequate account of persons, which Moran insists must be a normative account of practical commitments, privileges, and responsibilities.

Moran's discussion of Moore's Paradox, and Wittgenstein's response to it, in chapter 3 is also subtle and illuminating. Moore pointed out that one cannot properly say, '*P*, but I don't believe it', or 'I believe *p*, but *p* is false'. These assertions are bizarre and infelicitous, but they seem to involve no formal contradictions, for there is generally no logical entailment from my believing something to its being true (with the possible exception of the *cogito*). One response to the paradox, in a roughly Wittgensteinian spirit (though it was not in fact Wittgenstein's considered position), is to insist that the assertions are in fact self-contradictory, for expressions like 'I believe …' are not meant to report or describe one's state of mind, but serve merely to present or express one's acceptance of the truth of the proposition that follows. In spite of its indicative mood, that is, the first-person phrase has no real psychological reference and so does not constitute self-knowledge, but merely avows an attitude.

That may be true sometimes, but it cannot be an adequate account of the first-person indicative use of such verbs generally. As Moran points out, such locutions often, even typically, function as both avowals and reports, for I often report my attitude precisely by avowing it. The expression of attitudes must play a constitutive role in self-knowledge, otherwise self-knowledge would not strictly speaking be expressible at all, only reportable or capable of third-person attribution. In that case, the only kind of knowledge I could have of my own mind would be the peculiarly alienated sort of knowledge I have when I merely describe my beliefs as I would the beliefs of someone else. But again, recognizing those beliefs as my own, rather than as simply occurring in my mind, would still require that I be committed to them, that I take them up

and express them as mine. And no amount of third-person attribution of beliefs to myself will force me to do that.

Indeed, Moran argues that no amount of psychological evidence available to anyone to the effect that I believe *p* will even count as evidence that I believe it, unless at some level I also endorse the belief. Consequently, no exclusively third-person construction of my beliefs can show those beliefs to be mine, since *that* fact can be established only by my endorsing them from the first-person position. An entirely third-person view of myself and my beliefs, one might say, would leave no room for the *self* to whom such attributed beliefs are to be attributed. Warranted third-person attribution of beliefs thus turns out to be conceptually dependent on their first-person avowal. This is a powerful point, at least with respect to beliefs one considers reflectively and deliberates about, or what we might call 'opinions'. It is less obviously true of other kinds of beliefs, beliefs I simply find myself with, for example perceptual beliefs. As I shall argue, Moran's argument, here as elsewhere, seems narrowly focused on specifically deliberative attitudes at the expense of our mundane prereflective forms of understanding.

What is most exciting about Moran's project, it seems to me, is his effort to direct philosophical attention back to our ordinary concept of persons while preserving crucial Wittgensteinian and Sartrean insights into the asymmetry of the first- and the third-person standpoints. Wittgenstein, for instance, saw deeply into the asymmetries between first- and third-person discourse – so deeply in fact that he arguably found himself unable to offer any plausible account of how our concept of personhood hangs together as a single concept, instead of simply falling apart into first-person expressions and third-person descriptions of experiences and attitudes.[4]

Sartre insists on a comparable asymmetry in his distinction between the self as 'transcendence' and the self as 'facticity', which is to say, as freely and consciously directed *toward* the world, and as a mere thing *in* the world. Our dual status as both subjects and objects leads Sartre to such disconcertingly paradoxical formulations as 'the being of *for-itself* [consciousness] is defined … as being what it is not and not being what it is', and 'We have to deal with human reality as a being which is what it is not and which is not what it is'.[5] More plausibly, Sartre maintains that 'character has distinct existence only in the capacity of an object of knowledge for the Other. Consciousness does not know its own character – unless in determining itself reflectively from the standpoint of another's point of view'.[6] My own character remains invisible to me when my consciousness is, as it must be, directed toward the world. The claim is by no means absurd, but it suggests that the only kind of knowledge I can have of my own character will be a peculiarly estranged sort of knowledge, a knowledge I must borrow from an alien standpoint that I can never fully inhabit.

Finally, although Moran does not mention it, Heidegger's account of

selfhood in *Being and Time* can be said to exhibit the same basic incoherence, perhaps even more egregiously. For in the course of interpreting Dasein as 'being-in-the-world', Heidegger in effect repudiates all self-interpretations mediated by others as improper, disowned, or inauthentic (*uneigentlich*) in the evaluative as opposed to the merely formal or structural sense of that term. By conflating propriety or authenticity in the structural sense of first-person immediacy with propriety or authenticity understood as a style of comportment possessing positive normative value, Heidegger simply bypasses the question of how and to what extent second- and third-person points of view might figure into an adequate understanding of ourselves. What we are left with is a rather lopsided portrait of authentic Dasein radically submerged in a headlong rush toward its own finitude on the one hand, and an impersonal social world of irredeemable superficiality, anonymity, and self-evasion on the other. Hardly an adequate account of the subtlety and complexity of self-knowledge and social existence.[7]

What much of the history of modern philosophy shows, then, is how tempting it can be to collapse the first- and third-person perspectives together in the name of a unified concept of self-consciousness or subjectivity. That concept proved to be incoherent because it failed to come to grips with the social conditions and the perspectival asymmetry of our self-understanding. What the insights of Wittgenstein, Heidegger, and Sartre demonstrate, by contrast, is how tempting it can be to throw up one's hands and forgo any unified concept of personhood as an impossible hybrid of two radically heterogeneous, perhaps tragically conflicted personal and impersonal points of view. But such revisionism fails in turn to capture the relative stability and coherence that our integrated self-conception actually manages to exhibit, in spite of its interpersonal complexity. Moran's book is philosophically exciting because of the steps it takes in exploring that elusive combination of asymmetry and interconnectedness that marks our concept of persons, and that is so familiar to us in our prephilosophical lives.

II

What I find unconvincing in the argument of *Authority and Estrangement* is, in a word, Moran's rationalism. The intellectualist leaning of the book is by no means extreme, but its presence is instructive. It makes itself felt in the heavy emphasis Moran places on belief, in contrast to other practical, in particular noncognitive, attitudes; in his narrowly deliberative construal of belief; in his reduction of first-person practical attitudes at large to those that are specifically reflective or deliberative; and in his suggestion that all such attitudes, precisely in order to be first-personal in the proper sense, must be responsive or answerable to reason.

Understood descriptively, that last claim strikes me as false. And yet I suspect that Moran intends it not as a descriptive observation, but rather as an a priori normative demand: that is, not that our attitudes *are in fact* sensitive to rational scrutiny, but that they *ought* to be. So construed, the demand strikes me as arbitrary, or at best self-fulfilling, at least up to a point. Whether and to what extent our attitudes must be responsive to the demands of rational reflection and deliberation, it seems to me, depends at least in part on the extent to which we decide to insist that they be. Such insistence, systematically pursued, can have an intellectualizing effect on our lives, at which point the claims of reason come to seem ubiquitous and inescapable, as opposed to merely contingent products of a specific culture of rational discourse. It is that assumption of the fatality and inevitability of the demands of reason that I want to question here, especially since Moran himself at times seems to recognize the limits of the intellectualism driving much of his discussion.

The rationalistic slant of Moran's argument first emerges in chapter 2 in his criticism of Charles Taylor's account of human beings as 'self-interpreting animals'.[8] Taylor argues that our self-interpretations themselves constitute us by informing and shaping our attitudes. My pride itself changes, for example, when I come to regard it as sinful. Moran seems attracted to this idea, but he worries that notions like description and interpretation have too little implication of *truth* to do the constitutive work Taylor has in mind for them. I can describe or interpret all kinds of things this way or that way, he suggests, without taking seriously the possibility that they may really *be* so. If self-interpretations or self-descriptions are to impinge on us and shape the contents of our attitudes, Moran argues, they can do so only by making definite truth claims about them. Beliefs make such truth claims, but 'new formulations, vocabularies, or languages for our emotions' (p. 52) do not. Mere interpretive coloring or shading of that kind, he insists, would just drift over the surface without effecting any new commitments in our attitudes.

Our desires too, Moran argues, precisely in order to be sensitive to our interpretations of them, must be cognitive attitudes, otherwise they would be unresponsive to our interpretative view of them. To understand how emotions could be subject to our own self-interpretations, he says, we must 'view them cognitively and as falling under rational criticism' (p. 54). Moran concedes that phobias do not as a matter of fact yield to such criticism, but he adds that we typically criticize such fears as irrational. Indeed we do, but must we? We might decide to accept a person's idiosyncratic fears as brute responses, after all, and leave it at that, provided they don't seriously disrupt his life (or ours). Insisting that the person be made to feel the pressure of reason as a matter of principle seems arbitrary at best, coercive at worst. At any rate, it appears that other interests will have to weigh in as well before we decide how the interests of reason ought to figure in our dealings with him. Compassion and humility

might require that we bracket the charge of irrationality as irrelevant or inappropriate in this case. To reply by insisting that the charge is nevertheless true, even if tactless, would be no better than insisting that primitive cultures really do have bad manners. What is at stake is the applicability of the categories in question, not the mere desirability or undesirability of forcing the issue.

According to Moran, then, genuinely constitutive self-interpretation would require that all the attitudes at play in the process be cognitive attitudes, for 'a new description of my emotion or belief is powerless to alter it unless I *believe* the description' (p. 55). Beliefs are what rationally compel changes in other beliefs. Merely adopting new descriptions, interpretations, formulations, or vocabularies, Moran thinks, implies a kind of arbitrariness, if not outright voluntarism.

This strikes me as wrong. Taylor would surely include beliefs in the class of attitudes shaping and shaped by our self-interpretative activity, but there is no reason to limit the process to, as Moran would prefer, 'something belief-like on both sides of the relation' (p. 54). Aspects, or modes of presentation in perceptual and psychological experience, are not themselves judgments, though they constantly inform and shape our thoughts. Consider caricatures and stereotypes: they do not advance explicitly articulated 'belief-like' claims, but instead simply present their objects as, say, sinister or grotesque or ridiculous. They do not assert anything; they just make things appear in a certain (favorable or unfavorable) light. Emotions typically function this way, too, coloring and presenting things under aspects, without necessarily conveying or even implying distinct judgments about their objects. It is obvious, for example, that I can feel ashamed without judging that I have done anything shameful, just as I can feel joy, indeed in such a way that the joy shapes my beliefs and desires, without judging anything to be particularly joyous. Boredom is not, nor does it entail, a belief that something is boring.

Plato offers a vivid example of this kind of subrational moral self-interpretation in Book IV the *Republic* when he describes Leontius lustfully desiring to look at a bunch of corpses, feeling ashamed of the appetite, but making no rational judgment that it would be wrong to look. The diminished role of rational judgment in the example is crucial for Plato's argument, which purports to demonstrate the possibility of inner psychic conflict between moral emotion (the 'spirited' part of the soul) and appetite, not *reason* and appetite. If the moral revulsion with which Leontius regarded his own desire were exclusively, or even largely, a rational judgment concerning its wrongness, we would be left with an argument individuating only two parts of the soul, not three.

Moran does not want to reduce emotion to judgment or belief, as some philosophers do. But, unlike Plato, he wants to restrict its role in our moral psychology to its 'belief-like' effects, specifically the contribution it makes to

rational reflection and deliberation. Thus, in the case of fear, 'so long as I am to understand my condition as *fear* of any kind, even irrational fear, I cannot fail to accept the relevance, the force of the deliberative question, "Is there anything to be feared here?" ' (p. 63). But again, depending on a whole host of other considerations, that deliberative question might be relevant, and it might not be. It is true that emotions like fear are intentionally directed at objects, so the question, '*What* am I afraid of?' will always be to the point. But intentionality and rationality are not the same thing, which is why that question is not the same as the question, '*Should* I be afraid?' Should I fear heights? When there is no danger of falling, as in a (well-constructed) glass elevator, the fear has no rational basis. But does it need one? Such fear, at least in my case, is not a fear of falling, exactly, but is motivated directly by the perception of the vertical drop itself. Should I not have that emotional response to that perception? Why not?

Not only is it wrong to say that emotions and desires must be responsive to reasons; the truth is that they must be, at least to some degree, *un*responsive and resistant to rational reflection. Their affective force, their peculiar experiential inertia, after all, requires that they not be immediately sensitive to rational considerations in the way beliefs and judgments are. Moran seems to acknowledge this when he writes, late in the book, that

> ceasing to be ashamed or to feel ashamed is not the same thing as ceasing to believe that one did something shameful. … There is a temporal dimension to the moral meaning of various attitudes that is difficult, perhaps impossible, to capture in the terms of criticism developed for the evaluation of beliefs as true or false, justified or unjustified. (p. 180)

This seems right, indeed it strikes me as an understatement, but it also seems to be in tension with Moran's insistence that emotions figure in our first-person ethical attitudes owing not just to their intentionality, but to their having (or perhaps needing) rational foundations. Moran's discussion of shame and forgiveness in chapter 5 is deep and subtle, but it is not clear that his position can accommodate the essentially nonrational character of those emotions. We often have reasons to forgive or not to forgive, but purely rational considerations can never capture the unique intelligibility of forgiveness. If they could, then forgiveness would simply coincide with judgments of blame and come either immediately or never. Imagine a saintly person who judges guilty but then instantly forgives, with none of the usual 'cooling-off' period in between. Assuming the forgiveness is sincere, we could call such a person uncommonly charitable, but not irrational or inconsistent.

Finally, I find it an odd feature of Moran's argument that he seems to want to reduce the entire first-person practical standpoint to the sphere of rational deliberation. He writes, 'When the articulation or interpretation of one's

emotional state plays a role in the actual *formation* of that state, this will be because the interpretation is part of a deliberative inquiry about how to feel, how to respond' (pp. 58–9). As a categorical assertion, this is false. Perhaps I do sometimes deliberate about, maybe I can even *decide,* how to feel. But there is another way of actively engaging in the formation of one's own emotions through an effort of articulation and self-interpretation, and that is by focusing, refining, and as it were *cultivating* one's feelings by expressing them wholeheartedly. The point is not to decide how I *ought* to feel, but to get clear about how I *do* feel by letting my emotions take shape and find a voice in what I say and do.

What I have in mind here is incidentally not the kind of third-person exercise in psychological self-manipulation that Moran rightly wants to distinguish from the engaged, first-person perspective in which one's attitudes are 'transparent' and one attends not to oneself, but to the world. I think Moran is right that occupying the first-person position by conforming to what he calls the 'Transparency Condition' (which is to say remaining focused on the world and not retreating to a purely observational or psychological view of oneself) is not just a peculiar capacity of human consciousness, but is instead a normative demand we make in accordance with the very concept of a person. Expressing our attitudes *as* our own directly in the face of the world, as opposed to merely attributing them to ourselves as psychological facts from an observer's point of view, is a requirement, not just a fact, of personhood.

But I don't know why we should conceive of that first-person perspective solely, or even primarily, in terms of rational reflection. Moran acknowledges that it is not as if 'one normally arrives at one's beliefs (let alone one's fears or regrets) through some explicit process of deliberation. Rather, what is essential … is that there is logical room for such a question' (p. 63). I think it is important to recognize how weak that claim is. To say that there is 'logical room' for deliberative questions in practical life, I take it, is just to say that one can raise them without threat of inconsistency. But that says nothing about the relevance or propriety of such questions. That A leaves logical room for B is not a good argument for *assimilating* A to B. The constraints of logic alone are too weak to support Moran's idea that first-person practical attitudes can be understood just by understanding the position we occupy in rational deliberation.

Indeed, even in the case of belief, where Moran's intellectualist inclinations seem most at home, the privilege he affords to deliberative inquiry seems out of place. Moran appeals here to Stuart Hampshire, who writes,

> I make up my mind, and decide, when I formulate my beliefs. I do not observe them. But there are countless thoughts that occur to me, and that pass through, or that linger, in my mind, and of these only a small minority constitute beliefs. The beliefs are those

thoughts which I endorse as true. I do not merely find them occurring or lingering: I decide in their favor.[9]

As I said earlier, this may be a plausible description of the kinds of beliefs we call 'opinions', which do indeed rest on considerations of some kind, implicitly or explicitly, and require something like a verdict or decision in their favor, in the face of competing alternatives. But the passage does not describe the direct and fully involuntary beliefs we acquire, for example, by perceiving how things are. I look out the window and see that it's raining, so that's what I believe. I don't 'decide' that it's raining. I don't consider two views of the matter – (a) that it's raining, (b) that it's not raining – and then endorse one conclusion over the other. My belief is not a conclusion at all, but a kind of inheritance. Some beliefs are born, some are made, and some are thrust upon us. Moran seems to regard first-person beliefs primarily as things made, rather than born or received. Even in the domain presumably most hospitable to his rationalism, then, Moran's concept of belief seems to put too much emphasis on the role of rational reflection.

Granted, Moran does acknowledge that, just as desires are not always '*formed* as the result of deliberation', so too 'very few of our beliefs about the world arrive as the conclusion of any explicit *theoretical* reasoning that we undertake. It is nonetheless essential to the category of belief that a belief is a *possible* conclusion of some theoretical reasoning' (p. 116). But again, this is an extremely weak claim. Beliefs are expressed in propositions, and of course countless different lines of reasoning can be pieced together around them in an ad hoc way, like scaffolding around a building already standing on its own. But some beliefs in fact occupy a kind of bedrock or 'hinge' position in our system of beliefs, so that it would be perverse to ask for the rational considerations from which they are supposed to follow.[10] And while Moran is right to distinguish between an attitude being a *product* of deliberation and merely being '*answerable*' to deliberative considerations (p. 63), at other times he seems to blur that distinction and talk as if the intentionality of an action just consists in the structure of one's reasoning about it. Hence,

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. (p. 126)

Does this imply that in the absence of *reasoning* I cannot know the *reason* I'm doing what I'm doing? But if my 'guiding reason' is something cognitively unarticulated, such as an emotion or a desire, then I can indeed know what I'm doing, and even why I'm doing it, though reason plays no positive role in either generating or justifying my action.

Moran's fascinating discussion of shame and forgiveness, it seems to me, points to another limit of reason in the social dimension of our moral

practices. He refers to an episode in George Eliot's *Middlemarch* in which the young Fred Vincy has lost some money lent to him by the family of Mary Garth, with whom he is in love. Eliot shrewdly observes that Fred's primary concern, at least until he speaks to Mary face to face, is with his own moral image and reputation in her eyes and her family's. He desperately wants her forgiveness, but although she loves him, she senses that by granting it she would merely be indulging his narcissism, so she neither blames nor forgives: 'my anger is of no use', she says.[11]

Moran notes in passing that it makes a crucial difference to Fred's regret that it is Mary's judgment, not just anyone's, that matters so much to him, and he quotes Eliot's moving reference to 'an antique personage' who lamented after the death of his dearest friend, 'The theatre of all my actions is fallen'.[12] As Moran observes, Fred's deliverance from remorse 'is something he cannot do for himself – forgiveness requiring a person genuinely other to oneself' (p. 190). Someone genuinely other, but not just anyone. Fred Vincy is not alone in investing his moral self-image in the estimation of someone he cares about and whose attitudes matter to him, in some ways more than his own. Indeed, I suspect that this kind of social bond is an essential aspect of moral understanding. Moran does not pursue the idea further, but its importance might suggest that the rational standards we invoke in our moral reasoning are themselves beholden to contingent feelings of respect and affection that give us an intuitive sense that there is a perspective on ourselves other than our own, and that we depend on it for our very coherence and integrity as persons. That other perspective would be, in an important sense, a *second*-person point of view, not a third, and the personal aspect of the self-other relation might pose a challenge to the very idea of wholly impersonal standards of morality and rationality.

Earlier in the book Moran approvingly quotes Barry Stroud, who writes, 'The question of what I should believe in the situation in which I find myself is not the same as the question of what anyone who fits a certain general description should believe, even if I fit that description and know that I do.'[13] Stroud's point, and Moran's too, I take it, is that I can never regard my own attitudes as simply given, from a third-person point of view, but must be willing to consider them afresh, if I am genuinely to consider and assess any one of them. But there is a more radical inference to draw from the predicament of the first person, namely, that the rational and moral standards to which I submit are in a sense ultimately *my own* and need not be the same for anyone else who happens to fit the generic description that fits me. Explicit double standards may be impossible to justify, but it's not clear that wholly impersonal norms of thought and action can ever do full justice to the uniquely personal position I find myself in when I confront the question, What should *I* do?[14]

Abstraction and intellectualism nonetheless remain permanent occupa-

tional hazards in philosophy. Since Socrates, philosophers have felt the (probably unavoidable) temptation to regard the human condition at once generically and through the foreshortening lenses of their own reflective self-image. In succumbing to that temptation, they have tended to read their own rational commitments back into the subrational conditions of life that made the cultivation and articulation of reason possible in the first place. They thus find themselves caught between describing those mundane conditions in overly impersonal and rationalized terms on the one hand, and pleading the cause of further rationalization and refinement in our attitudes on the other. But those descriptions and those pleas must stand or fall on their own separate merits: we must try to describe our practices in all their rational and nonrational aspects as faithfully as possible, whether or not they are as they should be, and then consider how reflection, deliberation, and justification would figure best in those practices, whether or not they do. What we should not do is blur the distinction by treating our own current normative commitments as reason to suppose that our lives are always already rationally organized in the way we think they ought to be.

In any event, even if Moran's intellectualism fails to carry conviction, that failure scarcely detracts from the power and subtlety of this highly original book. *Authority and Estrangement* crosses conventional subdisciplinary boundaries in a fresh and exciting way and in so doing takes important steps in bringing personhood and self-knowledge back to the center of our philosophical concerns, and for that we can be grateful.

NOTES

1 See Descartes's reply to Mersenne, *The Philosophical Writings of Descartes, Vol. II,* trans. J. Cottingham et al. (Cambridge: Cambridge University Press, 1984), p. 100; AT VII 140.
2 Ryle, *The Concept of Mind* (New York: Harper & Row, 1949), p. 19.
3 Strawson, *Individuals: An Essay in Descriptive Metaphysics* (London: Methuen, 1959), ch. 3. Moran does not mention Strawson, but he comes close to Strawson's view, it seems to me, when he writes, 'the capacity for specifically first-person awareness of one's state of mind is necessarily tied up with *being* a subject of mental states in the first place' (p. 108). There is an essential connection, that is, between *having* an attitude at all and acknowledging oneself as its owner, which Moran in turn identifies with being in a position to *express* the attitude as one's own.
4 Hans Sluga makes this point, but notes that Wittgenstein also on occasion acknowledged the essential interconnectedness of subjective and objective uses of the first person. For example, Wittgenstein writes, 'The word "I" does not mean the same as "L.W." even if I am L.W. … But that doesn't mean: that "L.W." and "I" mean different things'. *The Blue and Brown Books* (New York: Harper & Row, 1958), p. 67. Similarly, 'But it is still false to say … I is a different person from L.W'. *Last Writings on the Philosophy of Psychology, Vol. II,* G. H. von Wright and H. Nyman, eds. (Oxford: Blackwell, 1992), p. 88. Quoted in Sluga, '"Whose House Is That?" Wittgenstein on the Self', in H. Sluga and D. G. Stern (eds), *The Cambridge Companion to Wittgenstein* (Cambridge: Cambridge University Press, 1996), pp. 346, 320.
5 Sartre, *Being and Nothingness: A Phenomenological Essay on Ontology.* Trans. H. Barnes (New York: Washington Square Press, 1966), pp. 28, 100.

6 *Being and Nothingness,* p. 457. Quoted in part by Moran, 192n.
7 For an elaboration of this argument, see ch. 6 of my *Heidegger's Analytic: Interpretation, Discourse, and Authenticity in "Being and Time"* (Cambridge: Cambridge University Press, 2003).
8 See Taylor, *Human Agency and Language: Philosophical Papers, Vol. 1* (Cambridge: Cambridge University Press, 1985).
9 Hampshire, *Freedom of the Individual* (London: Chatto and Windus, 1965), pp. 97–8. Quoted by Moran, pp. 113–14.
10 Hence Wittgenstein: 'At the foundation of well-founded beliefs lies belief that is not founded', in G. E. M. Anscombe and G. H. von Wright (eds) *On Certainty* (New York: Harper & Row, 1972), §253. For the 'hinge' metaphor, see §655.
11 Eliot, *Middlemarch: A Norton Critical Edition,* 2nd ed., G. G. Hornback (ed.) (New York: Norton, 2000), Book III, ch. 25. Quoted by Moran, p. 190.
12 *Middlemarch,* Book III: ch. 24. Moran, 189n.
13 Stroud, 'Practical Reasoning', in E. Ullmann-Margalit (ed.) *Reasoning Practically* (Oxford: Oxford University Press, 2000), p. 33. Quoted by Moran, 134n.
14 As Bernard Williams observes, 'Practical thought is radically first-personal. It must ask and answer the question "what shall *I* do?" Yet under Socratic reflection we seem to be driven to generalize the *I* and even to adopt, from the force of reflection alone, an ethical perspective'. *Ethics and the Limits of Philosophy* (Cambridge, MA: Harvard University Press, 1985), p. 21. Elsewhere Williams doubts whether the *I* can every be fully generalized in that way. See 'Moral Luck', in *Moral Luck: Philosophical Papers 1973–1980* (Cambridge: Cambridge University Press, 1981).

Received 4 June 2003

Taylor Carman, Department of Philosophy, Barnard College, Columbia University, 3009 Broadway, New York NY 10027, USA. E-mail: tcarman@barnard.edu