Two Conceptions of Coherence Methods in Ethics

Author(s): Michael R. DePaul

Source: *Mind*, Oct., 1987, New Series, Vol. 96, No. 384 (Oct., 1987), pp. 463-481

Published by: Oxford University Press on behalf of the Mind Association

Stable URL: http://www.jstor.com/stable/2253843

Two Conceptions of Coherence Methods in Ethics

MICHAEL R. DEPAUL

The method of reflective equilibrium defended first by John Rawls, and later by Norman Daniels, is the most widely discussed, most carefully specified version of a coherence approach to moral enquiry.[1] However, in spite of the detailed work that has been done on it, even this coherence method remains open to at least two distinct interpretations which differ with respect to the kinds of revision of pre-philosophical beliefs the method allows. Simply put, on a conservative conception of the method a person will revise moral or philosophical judgements where it is necessary to eliminate incoherencies latent in his belief system, with the particulars of these revisions being dictated by the person's initial commitment to the relevant propositions. A more radical understanding of the method recognizes revisions that go beyond what is needed to make a person's initial system of moral and philosophical beliefs coherent, allowing a person to alter her degree of commitment to propositions in ways that are not dictated by the logical or evidential relations between these propositions and the other propositions the person accepts or rejects. On the more radical interpretation, then, a person can simply change his mind about a proposition, while on the conservative understanding any change in belief or degree of commitment must be grounded somehow in elements of the person's initial system of beliefs.

The conservative conception has dominated recent discussions of moral epistemology and methodology, indeed, the radical conception of a coherence method has not even been recognized as a possibility. Yet it is only the radical conception that has a chance of being an adequate method of moral enquiry. In order to defend this claim I must first formulate more precisely the radical and conservative conceptions of a coherence method. I will then be in a position to support the radical conception through a consideration of the implications the two conceptions have for the epistemology of our moral beliefs and reflection on the way in which our moral views commonly develop.

[1] Reflective equilibrium was first described by John Rawls in *A Theory of Justice*, Cambridge, Mass., Harvard University Press, 1971, and 'The Independence of Moral Theory', *Proceedings and Addresses of the American Philosophical Association*, 1974/75, pp. 5–22. More detailed descriptions and a defence of this method can be found in Norman Daniels, 'Reflective Equilibrium and Archimedean Points', *Canadian Journal of Philosophy*, 1980, pp. 83–103, and 'Wide Reflective Equilibrium and Theory Acceptance in Ethics', *Journal of Philosophy*, 1979, pp. 256–82.

The two conceptions of reflective equilibrium agree to a basic outline of the method for moral theory construction. A person has little choice but to begin where she is, with her moral beliefs held with various degrees of confidence. A person is unlikely to be completely happy with all of these initial moral judgements, however, for after a little reflection most persons will grant that not all of their initial judgements were made in optimal circumstances, that is, in circumstances the person would take to be conducive to making an accurate judgement. Thus, for example, some of a person's initial moral judgements are likely to have been formed hastily or under the influence of some sort of bias, while the person might hold others with very little confidence. A person must therefore think over her initial moral judgements with an eye to eliminating those that are obviously epistemically sub-par. The remaining beliefs, still with degrees of commitment attached, are the person's considered moral judgements.[2]

In the next phase of reflective equilibrium a set of general principles which explicate[3] these judgements is formulated. In attempting to construct this explication neither the theory nor the considered moral judgements are granted a privileged status. A person must decide what to revise on a case by case basis by appeal to her degree of commitment to the propositions involved and the relations between these propositions and the other propositions she accepts or rejects. For example, if by slightly altering her considered judgements a person could accept a much simpler moral theory than the one which explicates the considered moral judgements with which she began, she may respond either by altering her considered judgements or by adopting the more complex theory. In all likelihood, a series of adjustments to both the considered moral judgements and the moral theory will be required to achieve a coherent system. When a person manages to bring such order to her moral beliefs she is said to be in a state of narrow reflective equilibrium.

The final phase of the method involves bringing into play the person's philosophical commitments that are not explicitly moral, for example, her views about the nature of persons, societies, or rational action. She does this

[2] I have here diverged somewhat from the ordinary description of reflective equilibrium. According to that description beliefs are filtered to obtain a person's considered moral judgements through epistemic principles which the person may or may not accept, either at the time of the filtering or in reflective equilibrium. This absolute appeal to epistemic principles is contrary to the spirit of coherentism. As I have described reflective equilibrium here, a person's initial moral judgements are filtered through her own epistemic principles. There are, quite obviously, obstacles in the way of saying just which principles these are. For example, they presumably will have to be something more than the epistemic principles the person happens to accept prior to any reflection. But I think it is quite clear that we will want a person's considered moral judgements to pass the muster of an internal rather than an external standard.

[3] A set of general moral principles is said to explicate a set of considered judgements when a person conscientiously applying the principles to the cases covered by the considered judgements would make the particular judgements which are either elements of the set of considered moral judgements or entailed by the general principles which are elements of this set. Cf. John Rawls, 'Outline of a Decision Procedure for Ethics', *Philosophical Review*, April 1951, pp. 184–6.

by considering alternatives to the moral theory she accepts in narrow equilibrium and attempting to formulate philosophical arguments that will decide among these competing moral theories. Once again, the objective is to revise beliefs in order to achieve the most coherent total system, without deferring to any of the component types of belief.[4] Thus, even if a person were to construct a philosophical argument against the moral theory accepted in narrow reflective equilibrium, say on the basis of her conception of a person, she would not necessarily reject the moral theory. She may choose to revise the conception of a person on which the philosophical argument rests. Wide reflective equilibrium is the point where a person's considered moral judgements, moral theory, and background philosophical beliefs form a coherent system.[5]

As I said above, both the conservative and the radical conceptions agree to this general description of the method of reflective equilibrium. The conceptions diverge with respect to the kind of revision involved in reaching wide reflective equilibrium.

According to the conservative conception of reflective equilibrium, a person is to achieve a coherent system of beliefs by formulating general moral principles that seem initially plausible and by investigating the logical and evidential relations among these principles, his considered moral judgements, and elements of his broader philosophical perspective. When the person uncovers a conflict[6] between two of his beliefs he will revise the one to which he is less strongly committed. The belief to be revised may be a considered moral judgement, a philosophical principle, or an element of his moral theory. More generally, when a person discovers that some number of his moral and philosophical beliefs are in conflict, he will select from the possible revisions which would yield a coherent system the one that maximizes his degree of initial commitment. In the simplest case this is the only type of revision required for a person to attain a point of wide reflective equilibrium. In such a case the person's equilibrium point is straightforwardly determined by his beliefs and degrees of commitment prior to

[4] By a type of belief I have in mind types that can be identified with reference to such factors as the form or content of the proposition believed or the methodological role the belief plays. It may be possible to identify types of beliefs that are immune to revision in other ways, for example, by appeal to a person's degree of commitment.

[5] Of course, in actual practice a person would not first strive to bring her moral beliefs into narrow reflective equilibrium, attempting to bring her non-moral beliefs into consideration only after she had reached a point of narrow equilibrium. A person aiming at wide reflective equilibrium would presumably be attempting to collate considered moral judgements, moral theory, and non-moral beliefs right from the start.

[6] I intend 'conflict' to be understood broadly. Thus, for example, a set of logically inconsistent propositions is in conflict, as is a set of propositions where all the members save one together provide strong evidence against that member. In cases of the latter sort, a person is under no obligation to resolve the conflict unless the conflict remains in the total set of his beliefs, since it is possible for a subset of a person's beliefs to provide evidence against a proposition when the total set of the person's beliefs provides strong evidence in its favour.

philosophical reflection. It will simply be the union of (i) the most comprehensive conflict free subset of his initial considered moral judgements and philosophical beliefs with (ii) a moral theory that explicates the moral judgements in this subset, which union preserves the highest total degree of initial commitment for the person.[7]

On this conception the method of reflective equilibrium approaches an algorithm for theory construction. The individual makes his contribution at the outset by specifying those propositions he believes and to what degree. But from this point on anyone with the relevant rational capacities could do the work of determining the equilibrium state; perhaps a machine would be best. The job consists of two tasks: (1) searching for conflicts among beliefs in the person's starting point and revising in a way that preserves the greatest degree of initial commitment, and (2) formulating a moral theory that 'completes the curve' begun by the considered moral judgements and background philosophical beliefs that remain after these corrections. Of course determining what propositions a person should believe in wide reflective equilibrium is at most half the battle. We would want to see the person actually come to hold this ideal system of beliefs, and bringing this about is likely to be no easy task.

Even on the conservative conception, determining an equilibrium point may not be quite so simple. In particular, two further sorts of revision of a person's initial beliefs might be needed. The first would be extremely rare, involving conflicts between or among beliefs to which the person is equally committed. It is perhaps not so very unlikely that a person initially would be equally committed to conflicting propositions, but this is not enough. To generate the kind of case I have in mind the person's commitments would have to remain equal after all the logical and evidential connections between the two propositions and his various other beliefs were taken into account. A case of such an all things considered tie among conflicting propositions would almost never arise, but if it did, the person would simply have to revise at least one of his commitments in a way that is not dictated by others of his commitments.

The second sort of revision involves propositions the person has never considered. Unless a person begins with a very broad range of beliefs, particularly moral, but also philosophical, some of the propositions he has not considered are likely to be theoretically significant. For example, a moral judgement about an unusual hypothetical situation might provide a good test for deciding between two moral theories that agree regarding all the propositions the person already believes. Or there might be metaphysical

[7] If we conceive of the process of achieving reflective equilibrium too strictly in terms of the probability calculus, we will run into trouble here. For as the number of beliefs with probabilities less than one increases the probability of their all being true decreases, and, therefore, in general it will not be possible to satisfy both the aim of maximizing degree of commitment and the aim of maximizing comprehensiveness.

questions, for example, about personal identity, that a person has never considered, but that might be relevant to certain of the person's moral judgements. In such cases, the person would have to consider the relevant propositions and make judgements about them for his equilibrium state to be decided. While it might seem that cases of the second sort occur all the time, for example, whenever a person is convinced by a counter-example, the appearance could be deceiving. For it is arguable that counter-examples do not function by eliciting new particular judgements which conflict with accepted moral principles, but by leading a person to use a more strongly held principle to decide a case about which a less strongly held, but more explicitly formulated, moral principle generates an incompatible judgement. And some might hold that in general, when persons make judgements about propositions they have never before considered, the judgements are based on the connections between the new propositions and the beliefs they already hold.

Still, because cases of these two types are possible, we cannot say that on the conservative interpretation a person's state of reflective equilibrium is completely determined by his starting point. The beliefs with which a person begins do, however, limit his equilibrium point considerably. On the conservative conception initial beliefs and degrees of commitment will not be altered unless such alteration is necessary to resolve a conflict, and the majority of these revisions will be decided by comparing the degree of initial commitment to the propositions involved.[8] The only other alteration of the initial beliefs will involve the addition of new beliefs where they are needed either because the initial set does not decide among competing moral theories, or because it does not contain any propositions about some important moral or philosophical issue.

Speaking loosely, I think it is fair to say that according to the conservative interpretation a person's beliefs in reflective equilibrium are determined by his convictions when he begins to reflect about morality, however well or poorly those convictions happen to be developed at that time. The conservative interpretation of reflective equilibrium is static, for it understands a person's point of reflective equilibrium as an ideal state which is virtually a function of the person's actual system of moral and philosophical beliefs at a given time. The significant point is that the only changes within this system envisioned by the conservative conception are those required to make the initial system of moral beliefs cohere. This version of the method can be seen as aiming at making a person's moral convictions at a given time explicit and precise. Given this conception those convictions determine the outcome of the method.

[8] Recall that conflicts can develop between moral and philosophical propositions. For example, a person who highly values theoretical simplicity might find certain of his moral judgements conflicting with this value if he were to realize that by altering these moral judgements he could accept a more simple moral theory.

The radical conception of reflective equilibrium allows for, and indeed expects, revisions of beliefs and degrees of belief that go beyond what is necessary to resolve conflicts in favour of more strongly held beliefs. According to both interpretations of reflective equilibrium the transition from narrow to wide reflective equilibrium requires considering alternative moral conceptions as well as philosophical arguments designed to decide among them. But only the radical understanding appreciates the full significance of this feature of the method. On the conservative interpretation, the consideration of alternative moral conceptions involves little more than having a person take her own non-moral philosophical beliefs into account and give some thought to the moral theories others have formulated and defended. The person's background philosophical beliefs provide another possible source of conflicts to be resolved, and the work of other moral theorists might contain a theory that fits better with her considered moral judgements and background philosophical beliefs than any theory she can devise on her own. According to the conservative conception, then, we have here only another opportunity for more of the same basic type of revision. The radical conception, on the other hand, sees the consideration of alternative moral theories as one significant point at which a person may be led to alter the structure of her moral beliefs in a way that is not determined by her previous degrees of belief in the propositions involved and the logical relations among them.

Let us consider an example of something that is not, I hope, terribly uncommon: moral conversion. The work of certain philosophers, for example, Nietzsche, Marx, or Sartre, is particularly apt to cause a person to experience such a conversion. We shall take as an example a person who reads Marx for the first time and comes away from the experience a committed Marxist. How are we to think of such a shift in views within the context of reflective equilibrium? Adopting the conservative interpretation, we might suppose that the convert simply found in Marxist principles of justice a moral theory that explicates her considered moral judgements and that coheres well with her broader philosophical views. She simply discovers in Marx's writings a moral theory that fits her initial belief system better than any theory she was able to formulate on her own. Of course if a person adopts Marxism in this way, any appearance of the person having experienced a moral conversion will be superficial. The conservative interpretation also allows for more complex scenarios, where there does seem to be a real moral conversion. For example, since Marxist theory offers an explanation of how our moral judgements are corrupted by various social and economic conditions, a person might begin by being convinced of this 'socio-economic' explanation of her moral beliefs. Movement towards a state of reflective equilibrium could then be driven by epistemological principles that the person holds constant. Moral judgements would be revised on the basis of these principles as the convert becomes aware of

Marxist reasons for doubting her initial moral judgements. There are, no doubt, other scenarios compatible with the conservative interpretation of reflective equilibrium as well. The distinguishing feature of these accounts is the identification of some strongly held belief or set of beliefs in the convert's starting point that is used to drive her to resolve conflicts within her belief system in a way that eventually leads to the acceptance of a Marxist theory in reflective equilibrium.

There are, however, moral conversions that cannot be understood in terms of the conservative conception of reflective equilibrium. When a person understands the whole Marxist perspective she may simply find that view of the world compelling in its own right. As a consequence she just changes her mind about various of her beliefs, now believing propositions that she previously would have considered unlikely to be true. There is no question here of the person making these alterations in order to resolve conflicts inherent in her initial system of moral and philosophical beliefs. We can suppose that system to have been coherent and thoroughly opposed to Marxism. Thus, it would simply be false that Marxism provides the best explication of the considered moral judgements the person made prior to reading Marx, and false that Marxism coheres well with the person's views about the nature of persons or society. On the basis of the beliefs the person had prior to reading Marx one would have had to expect the person to find Marxism a repugnant doctrine. But instead we find the person making different moral judgements and feeling as though scales have fallen from her eyes so that she can at last make correct judgements regarding morality. Of course, the Marxist theory the person now accepts explains why she used to be committed to a system of bourgeois considered moral judgements, but unlike the case described above, this explanation was not the catalyst for the conversion. The person did not first accept this explanation of her moral judgements because it cohered with various elements of her initial system of beliefs and then base revisions of her considered moral judgements and moral theory on this explanation in conjunction with epistemological principles.

In a truly radical moral conversion changes in belief or degree of belief are not dictated by some strongly held considered moral judgement or background philosophical belief. The person simply abandons at least some of her old considered moral judgements or background philosophical commitments and adopts a new set of commitments. The distinctive feature of the radical conception of reflective equilibrium is that it allows for the kind of *discontinuous* revision of belief which is the distinguishing feature of a radical moral conversion. On the radical understanding of reflective equilibrium a person can just change her mind about something. There may be no way in which the person's new commitments are determined by the old nor any way to see them after the fact as the result of appreciating and resolving conflicts inherent in the person's considered moral judgements

and philosophical views. The radical conception understands that philosophical reflection about morality, in some cases involving no more than reading the moral views of other philosophers, can sometimes lead a person to abandon her old considered moral judgements and background philosophical commitments and adopt a largely new set of convictions. On the conservative understanding when a person changes her mind about a belief it is always because she has come to see that there is a conflict among certain of her beliefs and she changes her mind about one thing in order to continue believing something else. For the work of a moral philosopher to sway a person, the philosopher must latch upon something the person already believes quite strongly and use this belief to force the revision of other beliefs.

Although philosophers do not tend to pay more than lip service to them, there are experiences other than the consideration of philosophical arguments which are likely to give rise to a reordering of a person's moral beliefs, and these other experiences have not, but should be, recognized as playing a significant and legitimate role in the development of a coherent moral view. Indeed, radical revisions set off by philosophy are certainly rare when compared with revisions initiated by these other experiences. I have in mind such things as living through a situation in which serious moral choices must be made, vividly imagining such circumstances (perhaps by way of literature, theatre, or films), coming into intimate contact with people who accept very different systems of value or suddenly understanding or appreciating such a system, listening to music, viewing paintings, or perhaps even mystical experiences and experiences of the divine or sublime. This list is obviously neither systematic nor complete. In addition it might give the false impression that I think only 'special' life experiences have the sort of effect I am talking about. But I hope my list is sufficient to get across the sorts of experience I have in mind.

I do not see how one might go about categorizing such experiences nor is it clear how the changes they effect in a person's moral beliefs should be evaluated. However, it is clear to me that there are such experiences, that such experiences bring about changes in the moral beliefs of those who have them, that these changes can be appropriate to the experience or not (in particular, it is sometimes rational to alter one's beliefs in the face of such experiences and it would often be irrational not to), and, finally, that having some experiences of this sort is just about a necessary condition for achieving a point of reflective equilibrium of which a person can justifiably be confident.[9] The radical conception of reflective equilibrium recognizes

[9] These last two claims obviously do not follow from anything I have said to this point. In the section below where I consider the rationality of a person's beliefs in reflective equilibrium I present my reasons for these claims. I say that having some experiences that are apt to cause discontinuous revisions of belief is *just about* a necessary condition for achieving a point of reflective equilibrium of which a person can justifiably be confident to allow the possibility of there being 'moral prodigies'.

these points and allows for extraordinary, discontinuous revisions of moral beliefs in addition to the types of revision recognized by the conservative conception.

My claim that persons sometimes revise their beliefs in ways that are not dictated by other things they believe is sure to be challenged.[10] When we change our minds about something we are generally able to provide an account of the change along the lines suggested by the conservative conception, and many people feel foolish when they cannot. We like to think of ourselves as rational beings, and it is tempting to think that if we are to be rational we must change our minds about things only when we have reasons for doing so, where here we think of a reason as something else we believe. The radical conception of reflective equilibrium seems to sanction a willy-nilly formation of beliefs which must be objectionable. So even if radical conversions occur, they surely are not the sort of thing for which a method of moral enquiry must find a place.

I cannot here respond to this battery of objections, but perhaps I can help my case by bringing into focus what I take myself to be claiming and what I think is required in order to deny my claim. We have been supposing that when a person takes up the task of constructing a moral theory there is a fairly large number of moral and philosophical propositions the person accepts or rejects to some degree or other. It is as if the person has gone through a list of propositions, considering each in turn and assigning a degree of commitment. It seems obvious to me that at a later time, after a person has had more experience and thought more deeply about things, the person might decide that she was mistaken about some of the assignments she made before she began to think philosophically about moral theory. And I see no reason for thinking that a person must connect up such a change with some other belief and degree of commitment on her original list. Surely many of the moral judgements we made with absolute certainty when we were adolescents are changed in this radical, discontinuous way—we come to look at things differently, decide that we made a mistake, and believe differently without feeling any need to justify our re-evaluation on the basis of something else that we believed when we were adolescents and continue to believe. The radical conception is intended to reflect the importance of this kind of change in belief. To adopt the conservative conception and to deny the possibility of discontinuous revisions of belief is, it seems to me, to suppose that when we begin moral enquiry we already possess as much of the truth about morality as we ever will. The harvest is over and moral enquiry amounts to no more than separating the chaff from the grain.

[10] I first suggested the distinction between the radical and conservative conceptions of reflective equilibrium in 'Coherence Methods in Ethics and the No Contact with Reality Objection'. When I read that paper at the Central Division APA meeting in May 1986 Norman Daniels, who was the commentator, maintained that all conversions could be explained in terms of the conservative conception. When I read an early version of the present paper at the Pacific Division APA meeting in March 1987 my commentator, Bernard Baumrin, made the same point.

The radical conception of reflective equilibrium is not static, as the conservative view is. Because the changes in a person's system of moral beliefs envisioned by this conception are neither determined, nor even constrained in any straightforward way, by pre-philosophical moral beliefs, it is not accurate to regard reflective equilibrium, radically conceived, as a method for articulating the moral conception a person has when she undertakes the construction of a moral theory. To be sure, the point of reflective equilibrium remains an ideal state on the radical conception, but it is not an ideal that is determined by the system of moral beliefs a person brings to the method. I suggest that it would be more accurate to say that the ideal state aimed at by the radical conception is determined by a person's faculty for moral judgement. According to this interpretation the method of reflective equilibrium does not involve mere algorithmic transformations on a set of beliefs. The method rather directs a person to let her moral faculty operate in conditions that are favourable to its functioning, conditions that will allow this faculty the richest inputs, ample opportunity to uncover and correct incoherencies in its output, the chance to interact with other judgemental faculties, and perhaps most importantly, the opportunity to mature.

Particularly if one thinks of the radical conception in the way suggested in the previous paragraph, one is likely to question my categorization of this method as a type of coherence method or version of wide reflective equilibrium. For my talk of a faculty of moral judgement invites one to think of moral perception or intuition rather than the kind of balancing of beliefs that reflective equilibrium is supposed to call to mind. As a result my radical coherence method may seem to be more of a version of foundationalism. I do not think it really matters very much whether we think of the radical method as a version of coherentism or foundationalism, as a further development of Rawlsian reflective equilibrium, or a similar but distinct method. The important question is of course whether there are deficiencies with reflective equilibrium, and whether the radical method develops reflective equilibrium in a natural way to avoid the deficiencies. None the less, I am inclined to think that the radical method is a type of coherence method as well as a sibling of wide reflective equilibrium. In the first place, unlike traditional sorts of foundationalism or intuitionism, the radical method does not grant any sort of proposition a privileged status. In addition, it recognizes the importance of the sorts of revisions characteristic of wide reflective equilibrium and incorporates all the features of that method of enquiry. It deviates from reflective equilibrium as traditionally conceived only in allowing initial degrees of commitment to be revised where this is not required to resolve a conflict and in requiring a person to seek out experiences in addition to those required by the conservative method. The latter deviation can be seen as nothing more than a way of ensuring that the enquirer has achieved a coherent equilibrium point which takes the most

comprehensive set of beliefs and experiences into account, and is therefore a more stable equilibrium than that attained by the conservative method. And the former deviation can be seen as moving even farther away from foundationalism than the conservative conception since it ensures that a person's initial beliefs and degrees of commitment are not playing a foundational role in the process of theory construction.

If I have said enough to get across what I have in mind by the radical and conservative conceptions of reflective equilibrium, it is time to say something about the significance of this distinction. In particular, I would like to suggest how the choice between these conceptions might affect the epistemology of moral belief.

It is now common among epistemologists to distinguish between two senses of rationality or justification. The one sense is a more internal notion, having to do with what a person can be praised or blamed for in her beliefs, with what she has an intellectual right to believe. The other notion is more objective, concerning the relation between the person's beliefs and truth. This objective notion is more plausibly thought to be involved in the distinction between knowledge and mere true belief.[11] To keep matters straight, I shall talk about the first notion in terms of *rationality*, and the second in terms of *justification*. The choice between the radical and the conservative conceptions of reflective equilibrium is relevant to our understanding both of the rationality of our moral beliefs and of the justification of these beliefs.

It will seem obvious to many that even conservatively interpreted reflective equilibrium will provide a method for attaining rational beliefs. For surely when a person has inconsistent beliefs these cannot be rational, and surely the way to correct this problem is to reject the belief of which one remains less confident after due consideration. And this is exactly what reflective equilibrium, even conservatively understood, directs one to do. Another reason for which we commonly criticize a person's beliefs concerns whether the person has surveyed the relevant alternatives to his position. Thus, if a person believes one explanation of a given occurrence when there is a readily available alternative explanation that coheres better with the person's overall view, but which the person has not bothered to think about out of sheer laziness, then we are inclined to criticize the person's belief as irrational. Here again it seems that reflective equilibrium leads one in the right direction, for it requires that a person consider alternative moral theories and relevant philosophical arguments in order to move from narrow to wide reflective equilibrium.

It seems to me, however, that the conservative conception does not take us far enough in either of these directions for us to be confident that a

[11] Cf. William P. Alston, 'Concepts of Epistemic Justification', *The Monist*, 1985.

person's beliefs in reflective equilibrium will be rational. It is entirely possible for a person initially to believe one proposition more strongly than another, but upon considering the fact that they conflict, become more strongly committed to the other independently of any connections between these propositions and other propositions the person believes. Depending upon the circumstances in which such a change of heart occurs, for example, whether the change is the result of a bump on the head, caprice, or careful reflection, I am inclined to think that it can be rational for the person to revise according to his *new* level of commitment.[12] However, since the conservative method aims at maximizing degree of initial commitment, it simply does not contain any provision for this sort of revision. In the situation described the person would either have to follow the directive of the conservative conception, that is, revise the belief with the lower degree of initial commitment, or start all over again, taking the beliefs with their new commitment levels as his initial moral judgements.

While the conservative conception of reflective equilibrium directs that one take alternative moral theories into consideration, it ignores other kinds of experience that might cause a person to alter his moral beliefs. I certainly would not want to say that whenever there is an experience that would lead to significant alterations in a person's moral view if the person were to have it, the person is not rational in those beliefs that would change in the face of this experience. After all, the experience of being brainwashed or having certain areas of the brain destroyed could lead a person to alter his moral beliefs. But typically a person will need to have had some experiences of the types that are apt to cause discontinuous revisions in a person's moral view before we should be confident of the rationality of his beliefs. An example will help establish the plausibility of this claim.

Suppose that a person does not see unspoiled natural areas or wild animals as good at all, being interested solely in the exploitation of resources for economic gain. Such a person clearly could incorporate his beliefs regarding the moral insignificance of the natural environment into a point of wide reflective equilibrium, conservatively attained. For we can suppose

[12] To be more exact, whether it will be rational to revise in terms of the new level of commitment will depend upon what the person believes about the circumstances in which he experienced the change of heart. Thus, if the person believes that he is less sure of a belief than he used to be because he is exhausted and not thinking clearly, then assuming the person has the usual sorts of beliefs about his reliability when in such a state, it will not be rational to revise in terms of the new commitment level. However, if the person has no reason to be suspicious of the circumstances in which he comes to feel less of a commitment to a belief, I am inclined to think not only that it could be rational of him to resolve the conflict in terms of his new level of commitment, but that it would be irrational of him to do otherwise (barring the interference of some other strongly held belief, for example, that his recent change of heart indicates that he has no settled opinion and in such circumstances is an unreliable judge). Frankly, it strikes me as absurd to think that when a person comes to see that two of his beliefs are in conflict, and upon considering the matter as carefully as possible the person feels much more confident of one of the two propositions, that is, he considers it much more likely to be true, it would be rational for the person to attempt to disbelieve this proposition because he used to be more strongly committed to the other (again assuming that no other beliefs interfere).

this person to have invested a fair amount of time studying questions concerning the development of natural resources. The person could understand the relevant facts, know about the effects of various development schemes, and be acquainted with the arguments offered by environmentalists. But knowing how the land will be changed by, for example, a strip mine or shopping mall would not force this person to alter his stand on development. For this knowledge would generate no conflict in the person, since he does not feel the slightest inclination to think that a virgin hardwood forest is somehow better in itself than a heap of mine tailings or that a richly inhabited marshland is better than acres of asphalt parking lots. The arguments offered by environmentalists do not bring any conflicts within the person's belief system to light either, since these arguments rely on the rejection of the anthropocentric value system which represents the person's most deeply held considered moral judgements and philosophical principles. Hence, the consistent set of considered moral judgements, moral theory, and philosophical beliefs with the greatest total degree of commitment for this person could well contain the belief that wilderness areas and wild plants and animals have no intrinsic value. However, I am not at all convinced that we should grant that this person's view regarding the value of wild things is rational. For suppose that the person has had no first-hand experience with wild places or things, that he has never even watched any nature films, and that he has not read any literature that seeks to convey an appreciation for the value of the natural world, for example, the poems of William Wordsworth or Walt Whitman. The supposition that the person's experience with the unspoiled environment is limited in this way makes it tempting to conclude that his beliefs about the moral value of natural things is irrational, but I think that this would be a bit hasty. For it might be that the person's 'data' are limited through no fault of his own, perhaps because he was reared in a purely urban setting among people who have never considered the possibility that the natural environment is valuable as anything other than something to be exploited. Or it might be that the person would retain his beliefs about the intrinsic value of the environment even if he were to have experiences that generally engender an appreciation of nature in persons, for example, strolling through a beautiful old forest, watching animals go about their own business in the wild or reading Henry David Thoreau. In either of these cases, I am willing to grant that the person's moral beliefs about the environment might be rational. However, we can suppose that the person lives in our society, where environmentalists are busy doing all they can to make available to people experiences which are apt to help them appreciate the natural world, that the person recognizes the way these experiences generally affect people, but purposely avoids these possible threats to his viewpoint, and finally, that if he were to have the experiences in question he would simply abandon his belief in the morality of wantonly exploiting the environment.

I think that I have described a sort of person and circumstance of belief where we would want to hold that the person's beliefs regarding the value of wildlife and wilderness are *irrational*.[13] I would like to be able to go on to argue that the radical conception of reflective equilibrium would lead one to avoid this sort of irrational belief since it directs that in addition to examining the logical and evidential relations among the beliefs one already holds and considering alternative moral conceptions and relevant philosophical arguments, that one seek experiences that have a significant impact on moral and philosophical judgement. However, I cannot derive this kind of direct support for the radical conception from reflection about the rationality of belief. One problem is that in situations similar to the one described above, indeed, in situations I am unable to distinguish in a relevant way from the one described above, the person's moral beliefs prior to the disruptive experience seem to be rational. In fact, it is clear that it is sometimes rational to avoid experiences that are likely to lead to a radical revision of one's moral views.

Consider the views of an Amishman. The Amish feel morally obliged to live a simple agrarian life, apart from the modern world, working in harmony with nature. We need not suppose that the Amishman has merely gone along with the training he received as a child and the pressures to conform exerted by his tightly knit family and religious community. We shall instead suppose that this man has thought a great deal about the rather odd way in which he lives as compared with the modern American society that surrounds him. He does not reject the use of various elements of modern technology on the basis of blind obedience to tradition or some kind of simple-minded notion that the Bible forbids the use of this technology. Rather, we suppose him to have a well worked out conception of the kind of life it is fitting for a person to live, in particular, the kind of relations a man

[13] I want to make it clear that I am not relying on any supposition that the relevant beliefs are false. Indeed, in so far as the person I have described takes wild things to have value only in virtue of human interests, his views lie in the mainstream. My point is that, true or false, this person's beliefs are irrational because he has negligently failed to seek out available experiences which would be relevant to his beliefs about the moral value of natural objects. From certain perspectives on rationality, in particular strongly internalist perspectives, my claim that the person's belief is irrational will seem plausible only because I supposed that the person is aware of the relevant experiences and their potential impact upon his beliefs. I am not sure that I agree with this perspective on rationality. But even if it is correct, I think that it is quite clear that if a person has undertaken to construct a moral theory, but has neglected to expose himself to relevant experiences that are available to him (even if the person is not aware of the experiences in the sense involved in determining rationality), then that person's moral beliefs are epistemically deficient in a way an adequate method of moral enquiry must avoid. The situation here is similar to one we might encounter in science. While we might not want to say that a scientist is irrational in believing a theory when published experimental results of which the scientist is completely unaware disconfirm it, it is clear that something is epistemically amiss with the scientist's belief, and that, whatever we decide to call it, the deficiency is one that a proper method of enquiry ought to help us avoid. I would say, therefore, that even if the externalist assumptions I have made about rationality are incorrect, and I am wrong to have stated my case in terms of the rationality of belief, the general point I am trying to make would survive with a minor reformulation of my presentation. I am indebted to Stefan Sencerz and Richard Feldman for discussions of this point.

should have with his family and neighbours and the kind of relation a man should have to nature and his work. It is this conception of the good life that grounds the Amishman's rejection of most modern conveniences and with them the life-style we would consider normal. In short, although he would not describe it in these terms, we are imagining that the Amishman has attained a point of reflective equilibrium, apparently conservatively understood, in which he holds a simple nineteenth-century agrarian life-style to be morally required of him. Now, the Amishman is certainly aware that there are many experiences which he has not had: experiences which we might take to represent some of the good things about the life-style that he has rejected. I have in mind such experiences as watching television, going to movies, reading newspapers, using modern labour-saving devices in one's daily life, and even such things as pursuing higher education. The Amishman is unlikely to have so high an opinion of his integrity that he would not recognize that were he to enjoy enough of these experiences he would no longer be committed to the moral view he presently accepts. But rather than thinking that he should expose himself to both the experiences constitutive of his way of life and also those constitutive of the alternative in order to be in a position to make a rational choice between them, the Amishman will feel no need to sample our life-style. He will think that the experiences he lacks would corrupt his judgement rather than inform it, and hence, he will avoid these experiences.[14]

Even though I do not know how to draw a principled distinction between the Amishman and the developer, I think that it is quite clear that the Amishman is not irrational in accepting his moral theory.[15] It might therefore appear that the radical and conservative conceptions of reflective equilibrium are on an equal footing. In certain circumstances one following

[14] Many of the issues touched upon here arose in a practical setting when states tried to force the Amish to send their children to public schools. The Supreme Court finally ruled in favour of the Amish in *Wisconsin v. Yoder*, 406 U.S. 205 (1972). Although First Amendment issues regarding religious freedom were fundamental to the legal decision, it is quite clear that the heart of the Amish position was that the experiences their children would have in public schools beyond the eighth grade would corrupt their children's moral judgement so that the children would no longer be capable of affirming characteristic Amish values. One argument used by the state to justify sending the children to public school was that they would be in a position to make a rational judgement regarding the type of life they wished to live only if they had the kind of experience of the normal American life-style which is typically attained through a high school education. For a discussion of the Amish stand on education see Albert N. Keim, ed., *Compulsory Education and the Amish: The Right Not to be Modern*, Boston, Beacon Press, 1975.

[15] Lest it seem that I am being too generous in attributing rationality to the dogmatic beliefs of a member of an extreme religious cult, it is worth noting that in the relevant respects the Amishman is not significantly different from any of us. For our purposes the only remarkable thing about him is that he has a settled opinion that certain experiences would be corrupting, and this is not at all remarkable. For surely we all recognize that there are experiences that would probably lead us radically to alter our system of values, yet for the most part we do not feel that we ought to go out and have these experiences. We recognize that they would corrupt our moral judgement so we avoid them, and this, I think, is the proper attitude to have. A person would have to have an absurdly high opinion of his intellectual fortitude before he could think that there are no experiences which would corrupt his judgement, and that he therefore should be open to all 'potentially' corrupting experiences.

the conservative method will not avoid holding irrational beliefs that the radical method would have avoided, but in other circumstances the radical conception would lead one into irrational beliefs that the conservative method would have avoided. So the radical conception might seem to offer us no special advantage in our efforts to obtain rational moral beliefs.

There is a second problem with the claim that a person who follows the radical method of reflective equilibrium will be led to hold a rational system of moral beliefs. A system of beliefs in reflective equilibrium, on either interpretation, has a certain sort of stability. According to the conservative interpretation the system of beliefs is internally stable in the sense that there is no subset of beliefs in the system such that the person would have to alter his overall system of beliefs if he were to become aware of the logical or evidential relations among the beliefs in the subset. The radical conception of reflective equilibrium seeks to add a sort of external stability. It seeks to produce a belief system that will remain stable in the face of certain kinds of experience. But even if we acknowledge that there are experiences that bear upon the rationality of a person's beliefs of the kind the radical conception recognizes, it is not clear that the stability the radical conception seeks will protect one from holding an irrational system of beliefs. For suppose that there are a number of different kinds of morally significant experience, and that once a person has had a sampling of these experiences his beliefs in reflective equilibrium will remain stable in the face of further experiences, but that the *order* in which the person has these experiences determines his beliefs in reflective equilibrium. In particular, we might suppose that the person's point of reflective equilibrium is primarily a function of his last morally significant experience. In this case it would seem that we should be suspicious of the person's beliefs in reflective equilibrium, even radically conceived. It is a pure accident that he has the moral beliefs he has, stable though they are. It would not seem that the person should just go on holding his moral beliefs as though they were the legitimate outcome of rational enquiry; he should recognize that these beliefs are nothing more than a by-product of his unique personal history.

Rather than conclude that the radical conception of reflective equilibrium is in the same position as the conservative conception with respect to the rationality of the system of beliefs it finally produces, I think that the two problems just described should be turned to the advantage of the radical conception. For both these problems, but particularly the former, are truly deep problems about the rationality of moral theory. Aristotle claimed that the young were not fit to do ethics—that they lack the necessary experience.[16] The radical conception of reflective equilibrium is sensitive to this point. When one begins to look at moral enquiry in the way this method suggests one confronts the problem of trying to determine which kinds of

[16] Of course their lack of experience was not the only, or even the primary, reason Aristotle gave for thinking they are unfit for ethical theory. Aristotle's primary reason for thinking the young are unfit for ethics was their inability to profit from the study by translating their conclusions into action.

experience a person should have before we can grant the rationality of the person's moral position. This is as it should be. It is to the credit of the radical conception that the problem arises, and it counts against the conservative conception that this kind of question does not arise when we look at moral enquiry by its lights.[17] A similar point should be made about the problem of an accidental external stability. This is a real problem. If it is the only kind of stability we can attain for our moral beliefs, then something is wrong with these beliefs. But it is no criticism of the radical conception of reflective equilibrium that it allows us to recognize this. It is, however, a serious flaw in the conservative conception that it focuses only upon eliminating internal instability, entirely ignoring problems of external stability.

I have suggested that the radical conception of reflective equilibrium is preferable to the conservative conception on two counts: (1) it is more true to the way philosophical reflection about morality actually proceeds in that it allows for the sorts of radical revisions that sometimes occur and recognizes the importance of various 'non-philosophical' experiences, and (2) it guides us away from kinds of irrational belief that the conservative conception does not help us avoid and highlights important questions about the proper role of experience in rational moral enquiry. But what of the justification of our beliefs in reflective equilibrium? It would be nice if it could be shown that our beliefs in reflective equilibrium are necessarily justified, but I do not think that this can be shown for either the conservative or the radical understanding. My aim is therefore more modest. I simply want to suggest how one might try to explain the fact that a person's beliefs in reflective equilibrium are justified if it ever turns out that such beliefs are justified.

If we have the conservative interpretation in mind, discerning the general shape of an account of the justification of a person's beliefs in reflective equilibrium takes no great exercise of imagination. This method offers a traditional picture of theory construction. A person begins with a large body of beliefs: moral, philosophical, and factual. She proceeds to tidy them up by eliminating such obvious epistemic messes as inconsistencies. She then tries to extend this body of beliefs by generalizing to a moral theory. Given that a person's initial beliefs and degrees of commitment play a foundational role in theory construction, it seems natural to assume that the epistemic status of the outcome of such a process will depend upon the status of the inputs.[18] And this assumption provides just the framework within which the discussion of reflective equilibrium has been conducted. Critics charge that

[17] For a more detailed discussion of the problems involved in trying to determine the proper role of experience in moral enquiry, including a more thorough discussion of examples analogous to the two presented above, see my 'Naivete and Corruption in Moral Inquiry', *Philosophy and Phenomenological Research*, forthcoming.

[18] I have argued against this natural assumption in 'Reflective Equilibrium and Foundationalism', *American Philosophical Quarterly*, January 1986, pp. 59–69.

the method is nothing more than a modern version of intuitionism.[19] It can lead to little more than a polishing up of the prejudices ossified in a person's considered judgements. Since these judgements determine the outcome, the method cannot be legitimate unless the reliability of these starting points can be demonstrated. Defenders have responded by reminding us of the distinction between wide and narrow equilibrium, stressing the fact that considered moral judgements do not play this decisive role since they can be revised in the face of philosophical arguments. They have granted that considered moral judgements need to be defended, but urged that this defence can only be provided within reflective equilibrium, not prior to the construction of a moral theory.[20]

It seems to me that this entire debate has presupposed (sometimes explicitly, sometimes implicitly) a kind of foundationalist, 'garbage in—garbage out' view of the epistemology of moral theory construction. This sort of idea is obviously inappropriate to the radical conception of reflective equilibrium, for according to it a person can simply drop the beliefs and degrees of belief that serve as the inputs to the method and adopt an unrelated system of beliefs. Thus, it is at least possible for a person to begin with a set of beliefs that are utter epistemic trash and still end up with beliefs that are at least true. If we are going to be able to tell a story in which a person following the radical method of reflective equilibrium ends up with justified beliefs, we need to think about the epistemology of moral theory construction rather differently.[21] I can do no more than suggest one way in which we might begin to do this.

The first step to providing a possible account of the justification of a person's beliefs in reflective equilibrium is to realize that on the radical conception this method does not look anything like a vehicle for transferring epistemic status. It is not a method of enquiry in the traditional sense which exclusively involves performing certain operations upon a set of propositions or beliefs. The distinctive feature of the radical method concerns the acquisition of new beliefs. We might therefore expect that, if it produces beliefs that are justified, it would do so in a way that is similar to the way in which other belief-producing processes that do not take beliefs as inputs yield justified beliefs. Thus, we should expect the epistemology of a person's beliefs in reflective equilibrium to look more like the epistemology of

[19] Cf. R. M. Hare, 'Rawls' Theory of Justice—I', *Phil. Quarterly*, April 1973, pp. 144–55, and Peter Singer, 'Sidgwick and Reflective Equilibrium', *The Monist*, July 1974, pp. 490–517.

[20] This response is offered by Norman Daniels in 'Wide Reflective Equilibrium and Theory Acceptance in Ethics', *Journal of Philosophy*, 1979, pp. 256–82.

[21] In saying that the foundationalist view of the epistemology of theory construction which has dominated the debate about reflective equilibrium is inadequate I do not imply that any foundationalist approach must be. The dominant conception runs into trouble not because it is foundationalist, but because of where it assumes the foundations must be located, that is, in a person's considered moral judgements. I allow that versions of foundationalism which locate the foundations for the justification of a person's beliefs in reflective equilibrium elsewhere may prove to be acceptable. Indeed, the account I shall suggest might be looked at as a type of foundationalism, albeit a version where the foundations of the justification of moral theory emerge in the process of theory construction rather than at the beginning where they are available to guide theorizing.

perception or memory than the epistemology of deduction or induction. To explain the justification of a person's beliefs in reflective equilibrium in this way we would have to conceive the activity of moral theory construction, or at least portions of this activity, as a sort of extended experience. The beliefs that one forms as a result of such an experience could then be seen as having their epistemic status as a result of their relation to features of this experience, for example, its reliability or intrinsic character, rather than on the basis of the logical or evidential relations among these beliefs and the beliefs one held when one began to reflect philosophically about morality.

Recognizing that particularly in the last section my remarks have been very sketchy, let me summarize what I take myself to have accomplished in this paper. I first distinguished two different conceptions of reflective equilibrium. These conceptions do not differ with respect to the sorts of propositions that are taken into consideration, as do wide and narrow reflective equilibrium. They differ rather with respect to the kinds of revisions that are allowed. The conservative interpretation essentially presents a procedure for resolving conflicts among beliefs on the basis of degree of initial commitment, while the radical conception endorses both revisions that do not rely on initial commitments as well as revisions that are not required in order to resolve conflicts inherent in a person's system of moral and philosophical beliefs. As a result, the radical conception has a different look from traditional methods of enquiry or theory construction. While the radical interpretation of reflective equilibrium diverges from paradigmatic methods of enquiry, I suggested that it accords more closely with the way we actually develop our moral views than the conservative conception. I then argued that the radical conception would lead a person to avoid types of irrationality in her system of moral beliefs to which the conservative conception is insensitive, and that the ways in which the radical conception might fail to yield a rational system of beliefs represent deep problems concerning the rationality of moral enquiry that must be addressed no matter what method of enquiry we adopt. In conclusion, I sketched what I consider a plausible way of accounting for the justification of a person's beliefs in reflective equilibrium, radically understood, although no claim was made that anyone's beliefs in reflective equilibrium would be justified. The point of this last section of the paper was merely to show how a person's moral beliefs might have a positive epistemic status.[22]

Department of Philosophy
University of Notre Dame
Notre Dame
IN 46556, USA

MICHAEL R. DEPAUL

[22] A number of people have provided me with such significant help in thinking about the issues raised in this paper that I must acknowledge my indebtedness to them. They are David Solomon, Richard Foley, Philip Quinn, Robert Audi, and Norman Daniels. I am also indebted to the Jesse H. Jones Faculty Research Fund at the University of Notre Dame for supporting the final stage of my work on this paper.