More on Williams on Ethical Knowledge and Reflection

A. W. Moore[1]

Abstract

This essay is concerned with Bernard Williams' contention in *Ethics and the Limits of Philosophy* that, in ethics, reflection can destroy knowledge. I attempt to defend this contention from the charge of incoherence. I do this by taking seriously the idea that ethical knowledge is knowledge from an ethical point of view. There nevertheless remains an issue about whether the contention is consistent with ideas elsewhere in Williams' own work, in particular with what he says about knowledge in *Descartes*. In an earlier essay I argued that it is not. In a subsequent essay I indicated that I had changed my mind and gave a more sympathetic account of Williams' contention. In this essay I set out the issues and say some more about my change of mind.

I am extremely grateful to Ana Falcato and Susana Cadilha for their kind invitation to contribute to this special issue of *Topoi*, designed to explore Bernard Williams' legacy on the twentieth anniversary of his death. The phrase that they have used as a title for this issue was one that they also used in their invitation: 'the formation of the moral point of view'. They subsequently expanded on this with a reference to what they called 'Bernard Williams' insights on how reason and emotions are brought together to form the ethical point of view'.

The metaphor of a point of view will be especially prominent in what follows, and I hope that this will make my contribution correspondingly fitting. Nevertheless, the first thing that I wish to do is to call into question the very phrase 'the ethical point of view'.[1] My concern is with the definite article. Williams certainly had an abiding interest in what it is to have *an* ethical point of view; but there is an important sense, on Williams' view, in which there is no such thing as *the* ethical point of view, any more than there is such a thing as *the* way in which one should live (see Williams (2006), pp. 20–21 and 52–53, and Ch. 9, passim). In fact many of the most interesting questions in moral philosophy are questions about how different ethical points of view relate to one another—in particular, about what relations of understanding and appraisal can exist between them.

Note that that there is not the same concern with respect to the phrase 'the moral point of view'. Williams follows other thinkers in distinguishing between the ethical and the moral (Williams (2006), p. 6). The ethical pertains to any way of answering the question of how one should live; the moral pertains to one particular way of answering the question of how one should live, in terms of a set of distinctive notions that centrally include voluntariness and obligation. This still leaves room for colossal variation within the moral, witness the fact that both Kant and any hardline act utilitarian would count as answering the question in this latter way. Even so, there is enough that unites answers of this kind to warrant talk of *the* moral point of view. Williams himself is certainly happy to talk in such terms (Williams (1981), p. 2). Indeed, it would not be much of an exaggeration to say that it is precisely because talk of the moral point of view has the warrant it has that Williams has the concerns about morality he has. The definite article nicely registers the importunacy and the arrogance which Williams finds in morality and to which he takes such exception.

But let us return to the issue of what relations of understanding and appraisal can exist between different ethical points of view. One of the most interesting cases is that in

---

✉ A. W. Moore
adrian.moore@st-hughs.ox.ac.uk

1   St Hugh's College, Oxford, OX2 6LE, UK

[1] I should add that I intend no criticism. The phrase served its purpose in Ana's and Susana's invitation; and there was enough in how they characterized their proposal to indicate their sensitivity to the point that I am about to make.

which, over time, one and the same person, or one and the same society, goes from occupying one ethical point of view to occupying another, allowing for retrospective assessment of the former from the latter. This connects with one of Williams' most striking and most notorious contentions: 'that, in ethics, *reflection can destroy knowledge*,' (Williams (2006), p. 148, emphasis in original).

What does Williams mean when he says that reflection can destroy knowledge? And how does this connect with the idea of an ethical point of view?

I have addressed these questions elsewhere (see e.g. Moore (1991), Moore (2003), and Moore (2019)). I shall do no more in this context than sketch some of the main features of what I have already written. But part of the reason why I wish to revisit these questions—I hope this will not appear too self-indulgent—is that some of what I have written indicates a significant change of mind on my part and I have previously said less about this change of mind than I would have liked to. My plan in this essay is to return to it from a slightly different angle, and then to say a little more about it.

First, it will be helpful to introduce some terminology. Let us call the case in which there is retrospective assessment by an individual or a society from one ethical point of view of another ethical point of view which the individual or the society used to occupy the Retrospective Case. And let us call the individual or the society in question the Subject, and the two points of view in question the Present Point of View and the Former Point of View. Then what Williams has in mind when he says that reflection can destroy knowledge is a possibility that arises in the Retrospective Case: the possibility that some of what the Subject knew from the Former Point of View is not available to be known from the Present Point of View and that the explanation for this involves the intervention of reflection. The metaphor of a point of view is already helpful here, because it puts us in mind of one very straightforward way in which knowledge can be lost. Think about how turning around through 180 ° and thereby coming to occupy a new point of view can deprive someone of knowledge they had before. Features of reality that were accessible from their previous point of view are no longer accessible from their current point of view, or not in the same direct way. A familiar children's party game illustrates this very simply. The participant begins by looking at a tray with a large number of objects on it and thereby knows what these objects are; the participant then turns around and is asked to recall what they could see on the tray; and they typically find that they have lost much of the knowledge that they previously had.

One fundamental issue concerning the ethical counterpart of this is how the Subject's reflecting on their knowledge—call the knowledge κ—can play an analogous rôle to turning around through 180°. Williams' account of this focuses on the concepts that are required to have κ. These critically include what he calls 'thick' concepts, such as the concept of blasphemy or the concept of chivalry, whose distinctive combination of description and evaluation helps to constitute the Former Point of View. K is knowledge to the effect that certain entities do or do not fall under the concepts in question. The reason why the Subject no longer has κ, indeed the reason why the Subject no longer occupies the Former Point of View, is that they no longer evaluate things in that way and hence are no longer able suitably to exercise these concepts. And the reason why it is *reflection* that brings about this transition is that reflection upsets the relevant evaluation, say by showing that it is associated with false beliefs or simply by drawing attention to alternatives. The Subject no longer has κ because they are no longer able suitably to think in those terms.

So far, one might think, so uncontentious. But a fundamental awkwardness arises when we ask in what relation the Subject now stands to κ. Twice in the previous paragraph I used the word 'suitably': the Subject is no longer able *suitably* to exercise the concepts in question and is no longer able *suitably* to think in those terms. This connects with an important distinction that we need to draw. Any thick concept can be grasped in one of two ways, the disengaged way and the engaged way. To grasp it in the disengaged way is to be able to recognize when the concept would correctly be applied, to be able to understand others when they apply it, and so forth. To grasp it in the engaged way is not only be able to do these things but also to share whatever evaluation it involves and, thereby, to be prepared to apply it oneself. (See further Moore (2023a, b), pp. 216–217.) My uses of 'suitably' in the previous paragraph were in effect gestures at the *engaged* way of grasping the concepts in question. The Subject still grasps them in the disengaged way. If they did not, their loss of κ would not be a result of reflecting, but rather of something more like forgetting. However, if the Subject still grasps the concepts in the disengaged way, then they can look back at κ and still recognize it as knowledge. The awkwardness comes in trying to reconcile this with the claim that they no longer have κ. How can the Subject know that κ is knowledge and yet not still be said to have κ? Williams' contention appears to be in danger of collapsing into incoherence.

One response to this worry, on Williams' behalf, would be the following. There is no incoherence in recognizing an item of knowledge as such without sharing it, provided that the recognition in question does not require direct insight into the content of the knowledge. Thus I can know that you know how to solve a Rubik's cube, say, because I have seen you do it often enough in the past, even though I have no idea how to do it myself. For that matter, I can know that I myself used to know how to solve a Rubik's cube, or indeed that I used to know what was on the tray to which I now have my back turned, even though I no longer have

that knowledge. Why should the Subject not similarly know that they used to have various items of knowledge about what is blasphemous, say, even though they no longer know anything of this sort?

But this response is unsatisfactory. We have conceded that the Subject grasps the relevant concepts in the disengaged way. And this is enough to give the Subject precisely what, according to this response, they must *not* have if Williams' contention is to avoid collapsing into incoherence: namely, direct insight into the content of κ. For, given that the Subject has such insight, they seem to be in a position to assert things of the form, 'I/We used to know that *p*', where what replaces '*p*' involves the relevant concepts. And even if this is not straightforwardly inconsistent with their asserting, 'I/We do not now know that *p*', then at the very least there seems to be an element of self-stultification in their asserting both of these things, akin to the element of self-stultification that G.E. Moore noted in someone's asserting something of the form, '*p*, but I do not believe that *p*' (see Sorensen (1988), Ch. 1). It seems to follow that the Subject must still have κ.

I have tried elsewhere to indicate a better response to the worry on Williams' behalf, which at the same time disrupts the line of thought in the previous paragraph (Moore (2003), § 3). This involves taking the metaphor of a point of view very seriously. Consider temporal points of view. We no longer occupy temporal points of view that we once did. So we are no longer in a position to know some of what we once knew from those points of view. For example, we are no longer in a position to know what we knew when we claimed, 'No-one has ever walked on the moon'. True, we are in a position to know that no-one *had then* ever walked on the moon. But that is different: that is not the present-tense knowledge that we had previously. Yet none of this precludes or is precluded by our having direct insight into the content of that present-tense knowledge. The point is this. We can manifest our insight by saying that what we knew then was—in the very terms that I just put it—*that no-one had then ever walked on the moon*. The change of tense precisely enables us to express our former knowledge even though we no longer have it. To express our former knowledge in this way it suffices that what we say now, when we use the expression 'no-one had then ever walked on the moon', and what we said then, when we used the expression 'no-one has ever walked on the moon', are, in Williams' own terms, 'tantamount to' each other (Williams (2006), p. 143).[2] Provided that the Subject can find something likewise

tantamount to the knowledgeable claims they used to make, from the Former Point of View, about what is blasphemous, the threat of incoherence in Williams' contention can be dispelled.

Admittedly, giving a precise account of what it is for one expression to be tantamount to another, in the sense involved here, would be a non-trivial matter. For instance, it does not suffice that the two expressions share their content, at least not on an undemanding conception of content. Suppose our claim 'No-one has ever walked on the moon' occurred exactly 1970 years after the birth of John the Baptist. Then, on an undemanding conception, the sentence 'No-one's walking on the moon occurs until at least 1970 years after the birth of John the Baptist' shares its content with what we claimed. Even so, it would be wrong for us now to assert, 'We knew then that no-one's walking on the moon occurs until at least 1970 years after the birth of John the Baptist'. We were not in a position to know any such thing.[3]

But we do not need a precise account of what it is for one expression to be tantamount to another for current purposes. For, even without such an account, we can see that the Subject is indeed able to find something that is tantamount to the knowledgeable claims they used to make, from the Former Point of View, about what is blasphemous. As Williams' own discussion of these issues indicates (Williams (2006), pp. 143–145), the very thing that looked as though it was creating trouble for his contention, namely the Subject's disengaged grasp of the concept of blasphemy, is the very thing that the Subject can exploit here. This is because the Subject's disengaged grasp of the concept allows them to exercise it vicariously. And this in turn allows them to say something that is clearly tantamount to what they said when they claimed, with respect to some work of art, say, 'This is blasphemous'. They can simply say, with respect to the same work of art, 'This is blasphemous'. True, they cannot say this *in propria persona*. But they *can* insert it in a context such as 'I/We would once have recognized that…'. They can do the latter because that involves nothing more than a vicarious exercise of the concept of blasphemy. It follows that the Subject is in a position to assert, without any inconsistency, and without any Moorean self-stultification, both, ' I/We

---

[2] It is worth noting that the construction described here varies from one language to another. In some languages, such as Russian, were we to report what we knew then we would have to retain the present tense to express our knowledge. But this does not gainsay the point I am making: if anything, it makes it more graphic.

[3] Here is another example—of a sort that has been much discussed—which illustrates essentially the same point. Suppose that Mr Meanour is suffering from amnesia after a car crash and has no idea who he is. And suppose that various clues have led him to say, knowingly, 'I'm wanted by the police'. Suppose, finally, that we want to report what he now knows. We had better not do this by asserting, 'He knows that Mr Meanour is wanted by the police'. His amnesia prevents him from knowing any such thing. For our expression of what he knows to be tantamount to what he himself has knowingly said, we need to use the third-person indirect reflexive counterpart of his own first-person pronoun. We need to assert, 'He knows that he is wanted by the police'. (Cf. Castañeda (1966) and Perry (1993)).

would once have acknowledged that this is blasphemous,' and, 'I/We no longer acknowledge that this is blasphemous'. What they are not in a position to assert, of course, is, 'This is blasphemous'.[4]

But now an opposed worry arises. If the Subject can exploit their disengaged grasp of the concept of blasphemy in this way—if this enables them to say something that is tantamount to the knowledgeable claims they used to make, from the Former Point of View, about what is blasphemous—then does it not also enable them to do something that Williams would insist cannot be done, namely, pick out what is blasphemous in neutral terms, that is in terms to which the Subject can accede even though they no longer evaluate things in that way?

Well, no. The Subject is not picking out what is blasphemous in 'neutral' terms. The Subject is picking out what is blasphemous by exercising the concept of blasphemy. It is just that they are exercising it vicariously. Williams is clear that the mere vicariousness of their exercise of the concept does not prevent it from doing its descriptive classificatory work (Williams (2006), pp. 141–142, and Williams (1995), pp. 205–206).

To take stock, then: I have tried to indicate why I think that Williams' contention that reflection can destroy knowledge is at least coherent. But it is a further question, of course, whether it is true. It is in connection with this latter question that I have changed my mind. In 'Can reflection destroy knowledge?' (Moore (1991)) I presented reasons for doubting Williams' contention. In a later publication, *Points of view* (Moore (1997)), I in effect retracted some of what I had said in that earlier piece (pp. 12 and 15–16). And in a yet later publication, 'Williams on ethics, knowledge, and reflection' (Moore (2003)), I registered this change of mind (n. 20) and proceeded to give a more sympathetic account of Williams' contention (§4). I now want to amplify a little on all of this.

In 'Can reflection destroy knowledge?' I raised the question whether, even if what Williams says about the Retrospective Case is internally consistent, it is nevertheless inconsistent with what he says in his book on Descartes (Williams (1978)). In that book Williams gives an argument that is based on something which he says appears 'basic to the notion of knowledge itself' (p. 65). 'If knowledge is what it claims to be', he writes, 'then it is knowledge of… what

is there *anyway*' (p. 64, emphasis in original).[5] It follows that, given any two items of knowledge κ_1 and κ_2 from two different points of view, for instance the Subject's one-time knowledge that some work of art is blasphemous and their later knowledge that the same work of art is, say, courageous, then 'there must be some coherent way of understanding why [κ_1 and κ_2] differ, and how they are related to one another' (p. 64). Let U be such a way of understanding κ_1 and κ_2. If U is to satisfy the demands placed on it by the fact that κ_1 and κ_2 are two items of *knowledge*, then it must do more than provide a coherent account of how they can be cohabitants of the same world. It must provide a coherent account of how they can be cohabitants of the same world *and count as items of knowledge*. It must represent them as being, not just from two different points of view *in* the same world, but from two different points of view *on* the same world. And, I urged in my essay, it cannot do this if it precludes their endorsement. For to preclude their endorsement is to preclude the very thing that allows them to be represented in this latter way. But if those who possess either κ_1 or κ_2 are capable of reflection, then not only must U be available, it must be available to them (Williams (1978), pp. 64–65). Hence, given that acquiring U involves standing back and reflecting on each of κ_1 and κ_2, it must be possible for those who possess either of these items of knowledge, if they stand back and reflect on it, to endorse it—which in turn means that it must be possible for those who possess either of these items of knowledge, if they stand back and reflect on it, still to possess it. Williams' contention that reflection can destroy knowledge is, I concluded, wrong.

One objection to my argument, which I did not consider at the time but which is immediately forthcoming from the discussion above, is that we can endorse an item of knowledge without ourselves possessing it, or even having the capacity to possess it. We can do this by asserting something that expresses it. Thus we can now endorse what we knew when we claimed, 'No-one has ever walked on the moon', by asserting, 'No-one had then ever walked on the moon', even though we can no longer know what we knew then.

This is certainly a legitimate objection to my argument. Even so, it is not a very damaging one. It does not prevent the argument from creating trouble for what Williams wants to say about the Retrospective Case. For there is no equivalent way of endorsing the Subject's one-time knowledge that some work of art is blasphemous without

---

[4] If there still appears to be some Moorean self-stultification here, reconsider the languages mentioned in note 2 above. If we are speaking one of those languages, then we are in a position to use a present-tense sentence to express knowledge that we once had, even though we are not in a position to assert that same sentence directly. For discussion of some closely related issues see Moore (2023b).

[5] Is there an issue about whether ethical knowledge 'is what it claims to be' (cf. Williams (2006), p. 199)? There are in fact compelling reasons why Williams had better not deny that it is. If he did, that would either significantly lessen the force of his claim that there is such a thing as ethical knowledge or significantly lessen the force of the argument he gives in his book on Descartes. For further discussion see Moore (2003), pp. 347–348.

actually acknowledging the work of art in question as blasphemous. Since Williams' contention that reflection can destroy knowledge applies to cases of this kind, the contention is still threatened by his own argument in the book on Descartes. For that matter, the very characterization of the Subject's various items of knowledge as items of knowledge is threatened by his argument in that book. For just as there is no endorsing the Subject's knowledge that some work of art is blasphemous without acknowledging the work of art in question as blasphemous, so too there is no endorsing the Subject's later knowledge that the same work of art is courageous without acknowledging it as courageous; and, if the details of the story are suitably filled in, there is no doing both of these things without, *per impossibile*, occupying two incompatible points of view.

I now think that there is a more telling objection to my argument. In my later essay 'Williams on ethics, knowledge, and reflection' I urged that the rôle assigned to endorsement in my argument should rather have been assigned to something less exacting which I am going to call 'indirect' endorsement (pp. 348–349).[6] Whereas endorsing an item of knowledge means committing to something that has the same content, if only on the undemanding conception of content to which I referred above, indirectly endorsing an item of knowledge means committing to something that *implies* the same content. Quite what this notion of implication comes to is a large and difficult question. In 'Williams on ethics, knowledge, and reflection' I suggested that we can indirectly endorse some item of ethical knowledge by saying enough about the cultural background of those who possess it, and about the particular circumstances in which they find themselves, to indicate how all their relevant beliefs not only involve faithful application of the thick concepts involved in them, but also succeed in tracking the truth. In fact I believe that we can, in principle, indirectly endorse some item of ethical knowledge in a way that is even more detached than that. I believe that we can do so by providing a purely physical description of those who possess the knowledge and of their circumstances (see Moore (1997), Ch. 4, §4). To be sure, the phrase 'in principle' is crucial here. It would be ludicrous to suggest that we can do this in any practical sense. But it is what we can do in principle that counts as far as what I have in mind is concerned. Clearly, far more remains to be said about this. In particular, a full account of what I have in mind would need, sooner or later, to invoke the notion of supervenience— and quite what *that* comes to

is itself a large and difficult question. For current purposes, however, only two things really matter.

First, as I urged in my later essay, the rôle assigned to endorsement in my original argument should rather have been assigned to indirect endorsement. In particular, all that is required for κ_1 and κ_2 to be represented as being from two different points of view on the same world, and hence all that U needs to make provision for, is their indirect endorsement. That is already enough to indicate how they are able to differ in the way they do while still counting as items of knowledge.

The second thing that matters is how detached indirect endorsement can be. In particular, while there may be no endorsing the Subject's knowledge that some work of art is blasphemous without acknowledging the work of art in question as blasphemous, the same is certainly not true of indirectly endorsing the Subject's knowledge. One way to appreciate this is to note that we could indirectly endorse the Subject's knowledge (albeit not in a way that would be sufficient for U's purposes) by doing something which the discussion earlier was precisely designed to show would not involve our acknowledging the work of art in question as blasphemous, namely saying, quite simply, that the Subject knew that the work of art in question was blasphemous.

These two things are together enough to scupper my original argument. They are therefore also enough, at least as far as that argument goes, to leave Williams' account intact. If there are any decisive objections to Williams' account, then they must go beyond anything that I offered there.

---

[6] Cf. Moore (1997), pp. 15–16, on which I am now drawing. But beware that I am using the expressions 'endorsement', 'indirect endorsement', and their cognates slightly differently here from how I use them there.

References

Castañeda H-N (1966) ' He': a study in the logic of self -consciousness. Ratio 8:130–157
Moore AW (1991) Can reflection destroy knowledge? Ratio (new Series) 4:97–107
Moore AW (1997) Points of view. Oxford University Press, Oxford
Moore AW (2003) Williams on ethics, knowledge, and reflection. Philosophy 78:337–354

Moore AW (2019) Realism and the absolute conception. In: Moore AW (ed) Language, world, and limits: essays in the philosophy of language and metaphysics. Oxford University Press, Oxford, pp 165–181

Moore AW (2023a) Maxims and thick ethical concepts. In: Moore AW (ed) The human a priori: essays on how we make sense in philosophy, ethics, and mathematics. Oxford University Press, Oxford, pp 210–225

Moore AW (2023b) Quasi- realism and relativism. In: Moore AW (ed) The human a priori: essays on how we make sense in philosophy, ethics, and mathematics. Oxford University Press, Oxford, pp 226–232

Perry J (1993) The problem of the essential indexical and other essays. Oxford University Press, Oxford

Sorensen RA (1988) Blindspots. Oxford University Press, Oxford

Williams B (1978) Descartes: the project of pure enquiry. Penguin, Harmondsworth

Williams B (1981) Persons, character and morality. In: Williams B (ed) Moral luck: philosophical papers 1973–1980. Cambridge University Press, Oxford, pp 1–19

Williams B (1995) Replies. In: Altham JEJ, Harrison R (eds) World, mind, and ethics: essays on the ethical philosophy of Bernard Williams. Cambridge University Press, Oxford

Williams B (2006) Ethics and the limits of philosophy. Routledge, London