11 The inevitability of inauthenticity

Bernard Williams and practical alienation

Nicholas Smyth

In *Ethics and the Limits of Philosophy (ELP)*, Bernard Williams declared that "ethical thought has no chance of being everything it seems." This cryptic and provocative remark went mostly unnoticed by his reviewers. Yet, as I will argue, it expresses a variety of practical skepticism which is central to Williams' thought and which deserves more attention than it receives. In what follows, I'll mainly discuss Williams' position itself, though I ultimately hope to show that any organized body of ethical thought must confront his particular brand of skepticism. Moreover, there is an important twist: I argue that Williams' own positive ethical philosophy is not exempt from this skeptical challenge, and I show that he did not clearly appreciate just how powerful the challenge is.

My first aim is exegetical. I want to produce the best account of just what Williams had in mind when he offered his striking thought. There has been a temptation to see Williams as simply echoing a variety of skepticism popularized by John Mackie, but I believe that it is of the utmost importance to see that Williams could not have meant to draw on Mackie's distinctly metaphysical brand of skepticism. Fundamentally, I will argue, Williams was telling us that any reflective, *normative* justification for our ethical dispositions will not match the justification that naturally comes along with such dispositions in ethical experience. The resulting post-reflective state should be described as a form of *practical alienation*, distinct from the metaphysical skepticism described by Mackie.

My second aim is more critical. I will argue that Williams' skepticism about ethical thought has devastating consequences for his own well-known arguments against various moral theories. The point will take some time to draw out, but it can be stated fairly simply. Williams wanted to claim that moral theory *alienates* agents from their projects and relationships, but his own skepticism about ethical thought implies that this sort of alienation is inevitable for reflective beings like us. I'll survey two possible solutions to the problem, but in the end I conclude that Williams cannot evade it. Furthermore, I'll finish with an argument designed to show that no ethical theory can avoid a confrontation with Williams' skeptical position.

These preliminaries aside, I'll now provide what I take to be the best interpretation of the claim in question. What does it mean to say that ethics cannot be everything it seems to be?

Practical alienation

Let's canvass a couple of possibilities, if only to dismiss them. One argument begins with an absolute distinction between factual and evaluative statements. Since, the argument goes, a complete description of the world and of our place in it will just be a list of factual statements, this description can contain no evaluative statements. Therefore, a complete factual understanding of the world and our place in it will fail to vindicate ethics.

A second line of thought, attributable to Mackie, begins with the claim that moral phenomenology presupposes the existence of entities which possess an objective "to-be-doneness," and concludes that morality suffers from massive presupposition failure since we have no reason to believe in such entities.³

It is important to distinguish these forms of skepticism about ethics from Williams' own. As we will see, the sort of reflective inquiry that generates his troubling conclusion is not *purely factual*, whatever that turns out to mean. Nor is it metaphysical inquiry into the nature of the entities which compose the world. Rather, it begins *within* ethics, so to speak. It is systematic, empirically informed investigation into the normative standing of our ethical beliefs and practices, performed in the light of what we already value. Since it operates mainly under the shadow of Mackie's challenge, contemporary metaethics is firmly focused on "placing" ethics in the natural world, but when Williams claims that ethics is not what it seems, he is not registering the failure of that project. Thus, Williams appears to be advocating for a distinct form of ethical skepticism, one which is not well represented in contemporary philosophy, and one which would remain pressing even if metaethicists managed to place ethical facts in the world. Perhaps, then, it deserves more attention than it has received.

That said, it is not entirely clear what Williams' skepticism amounts to, and I'll now proceed to offer a reading of the relevant texts. We should begin with his own elaboration on the claim in question:

The hope for truthfulness, next, is essentially that ethical thought should stand up to reflection, and that its institutions and practices should be capable of becoming transparent. I have tried to say why ethical thought has no chance of being everything it seems. Even if ethical thought had a foundation in determinate conceptions of well-being, the consequences of that could lie only in justifying a disposition to accept certain ethical statements, rather than in showing, directly, the truth of those statements: but this is not how it would naturally appear to those who accepted them.⁴

Let's tease apart and clarify the various thoughts involved in these remarks. We begin with reflective inquiry into the normative status of our ethical beliefs and practices. We move to a best-case scenario, where a true moral theory is actually discovered, or where a determinate foundation for ethical beliefs and practices is successfully defended. We discover that this reflective vindication of ethics is out of step with how it will "naturally appear" to ethical agents. This, we are told, is because this story merely vindicates the *disposition to have* those beliefs, and not the beliefs themselves.

This is more than a little puzzling. What does it mean to say that ethical thought "naturally appears" a certain way to people? Moreover, how exactly *does* it "naturally appear," and why is this appearance out of step with the justification provided by reflective inquiry?

In the following sections, I will argue that Williams' position contains three elements. First, a distinction between the "inside" and "outside" perspectives in ethics, second, a corresponding distinction between direct and indirect justification, and third, the idea of a personal project or social practice which is undermined by conflicts between the deliverances of the two perspectives. These ideas, taken together, guarantee that ethics cannot be everything that it seems to be.

"Inside" and "outside" perspectives

Reflective inquiry into our beliefs and practices involves adopting a perspective that Williams termed the "outside" perspective. It is important to distinguish this perspective from Sidgwick's "point of view of the universe," a perspective Williams consistently derided as a philosopher's fiction. Williams' outside perspective is not supposed to involve examining our ethical beliefs and practices sub specie aeternitatis, nor is it supposed to suspend the influence of the dispositions and values which constitute our practical identities.

That said, the idea of an "inside" or an "outside" perspective is not at all clear. Like Thomas Nagel's related distinction between subjective and objective perspectives, it can be accused of being *purely* metaphorical.⁵ However, it is clear that Williams believed the distinction to be of crucial importance, and as such, it is surprising that so little scholarly attention has been paid to it.⁶ In what follows, I'll try to remedy that situation.

Any philosopher who speaks of "standing back" from one's ethical dispositions owes us a psychologically realistic account of just what is going on when an agent performs this maneuver. Indeed, it is somewhat unfortunate that the metaphor of "standing back" is far more often invoked than explained in contemporary moral philosophy. Williams' opposition to Sidgwick's somewhat extravagant portrayal leaves him with a problem: that of explaining how an agent can "stand back" from her dispositions while nonetheless continuing to inhabit a practical perspective which is constituted by those same dispositions.

Here is what Williams should say. In evaluating a particular disposition "from the outside," we *simulate* the practical reasoning of an agent who does not have the disposition. We try to justify our possession of the disposition to an imaginary person who has some of our other dispositions but who doesn't have the particular one under scrutiny. Of course, we don't consciously imagine an actual conversation, rather, a subpersonal process carries out a kind of conversation between two practical reasoners, one with the disposition and the other without it. This simulational mechanism is familiar from contemporary discussions of "mindreading," of our ability to predict and interpret the actions of other persons. In his own discussion of action-interpretation, Williams explicitly endorsed this model, now most closely associated with Alvin Goldman.⁷

My suggestion is that he could draw upon features of this mechanism itself to explain what the outside perspective amounts to. Williams can say that the "outside" perspective is so named because it involves addressing the perspective of a simulated other. Plausibly, we learn how simulate in this way via our interactions with actual others whose evaluative perspectives differ from our own.⁸ Notice that the model itself implies something that is absolutely crucial to Williams' anti-Sidgwickian stance, namely, that in adopting the outside perspective we do not leave all of our evaluative dispositions behind.⁹ It is still *a particular person* who performs the "standing back" maneuver.

Yet, this model implies that the vindication provided by the outside perspective must be *indirect* in a sense I will now define.

Direct and indirect vindications

The next key to Williams' distinction lies in the *kinds* of justification provided by the two perspectives. In the quoted passage from the postscript to *ELP*, Williams' use of the word "directly" provides an interpretive key. For some proposition P, he contrasts an account which "justifi[es] a disposition to accept" P with an account that "shows, *directly*" the truth of P. Thus, imagine a person who judges that some friend is worthy of his loyalty. According to Williams, the "outside" perspective on this judgment cannot deliver the conclusion that the person *really deserves* that loyalty. Rather, the best we can hope for, from this perspective, is an account which shows that the *making* of this judgment is somehow valuable by the lights of some other set of values.

Something like this is suggested by what has just been said about "out-side" simulation. Consider again my loyalty to a friend. If the outside perspective on that disposition involves simulating the practical perspective of an agent who is not loyal to my friend, then this (imagined) other cannot possess any of the positive evaluations, emotional attachments, and historical experiences that help to constitute the disposition itself (my loyalty). But they are therefore not going to be swayed by appeals to those very

factors, which might serve to directly vindicate the thought that my particular friend is worthy of my loyalty. So, I am forced to relate my having the disposition to distinct things that the imagined other might value. This will be an indirect vindication, since it will explain the value of the judgment implicit in my loyalty not by citing its truth, but by relating it to some other set of values which is promoted or honored by my disposition to be loyal.¹⁰

This justificatory mismatch must be a key part of the story since Williams is clear in the quoted passage that the distinction between direct and indirect justification is key to his skepticism. Yet, I want to briefly argue that if this were all that Williams had meant to say, then his position would be uninteresting and very probably false.

Williams may have been tempted to say that ethics "naturally" appears to agents under the guise of evaluative judgments which appear to simply be true without qualification, but this claim would surely rely on an overly simplistic picture of ethical experience. After all, ethical agents combine direct and indirect thoughts as a matter of routine practice. In *Shame and Necessity*, Williams himself provides the example of Ajax, who combines the simple thought that he "must" commit suicide with a series of higher-order thoughts about shame, about his relations to others, and about his status as a warrior. To take another example, we do not need to be Rule-Utilitarians to see that "what if everyone acted this way?" is a basic ethical thought, as "natural" a part of ethical experience as any. Yet, the thought is manifestly a search for an indirect vindication of a first-order commitment, a vindication conducted in terms of some other, distinct set of values.

The outside perspective is no less "natural" than the inside one, and its appearances have as good a claim to be included among *the* ethical appearances. Why should a theoretical story—which merely tells us *how* to think when we occupy this wholly natural "outside" perspective—alienate us from our projects? And given the ubiquity of indirect reflection in ethical life, why should the mere distinction between direct and indirect justification deliver the conclusion that ethics is something other than it seems to be?

Though I believe we are now close to the correct reading, Williams cannot *just* have meant to highlight the fact that the "outside" perspective gives us indirect vindications of our beliefs and practices since those justifications are as much a part of ordinary ethical practice as anything else. There must be more to Williams' pessimism, and to discover what it is, we need to look at a distinct sort of practical situation which consistently fascinated and troubled him.

Substantive conflict

Williams certainly meant, in the quoted passages, to highlight the distinction between a direct vindication and an indirect vindication of some

ethical judgment. But as we have seen, he could not have meant to suggest that the "outside" perspective is something that philosophers invented and imposed upon first-order practice. Rather, he was also concerned to point out that the deliverances of the two modes can produce a certain psychological conflict in reflective agents. In other words, the difficulty is not due to *formal* differences between direct and indirect justification, it is due to tensions within the actual *content* of the justifications provided.

This idea is already implicit in the famous case of Jim and the villagers, where Jim's first-order commitment to nonviolence is in question. An indirect utilitarian is happy to say that this commitment is a good commitment to have, and for obvious reasons: it tends to maximize the good. Yet, Jim's situation is precisely one in which his disposition will *prevent* him from maximizing the good since a violent act will save many lives. Thus, his natural sense that violence is to be avoided conflicts with the thought that in this particular type of case, violence can maximize the good.

In his final book, *Truth and Truthfulness* (TT), Williams summarized the consequences of this situation in the following way:

The trouble is that an agent... has no thought to fall back on except that it is Utilitarianly valuable that he should have this disposition, and *this leaves no content to the disposition*: he has no thoughts with which to counter the consideration that some alternative action in this situation is the one that has the best Utilitarian consequences.¹⁴

If the outside perspective on Jim's disposition reveals that it is valuable because it generally produces good consequences, and if Jim accepts this, then there is a very real sense in which he no longer has the disposition in question. A settled disposition to avoid violence simply *is* the instinctive refusal to commit acts of violence; once it is regulated by reflection on general consequences, it loses the sense of "practical necessity" which, Williams claimed, accompanies our most basic ethical dispositions.¹⁵ This is what he meant when he described such dispositions as having "momentum," or, elsewhere, "a certain depth or thickness": phenomenologically, they appear as convictions that a certain behavior must or must not be performed.¹⁶

Williams is telling us that it is a deep fact about human beings that a gap must always open up between the two perspectives just described. Ordinary human commitments can be psychologically undermined if the outside, reflective mode is given priority, as Jim's commitment to nonviolence may be destroyed by his wholehearted acceptance of indirect utilitarianism. This is reflectively induced practical alienation in a nutshell.

At the social level, there are analogous dangers of institutional collapse in the face of external reflection. Edward Craig gives the useful example of a group of impoverished and endangered people who band together and enthrone a Hobbesian sovereign. Their shared, explicit justification for enacting this contract lies in the mutual security of all participants. But, for very ordinary reasons, a troublesome situation arises:

Then a rebellion breaks out and I am ordered into action to put it down, whereupon it becomes vitally important that I should have acquired a loyalty to the sovereign that is not simply a matter of my enthusiasm for the function for which he was enthroned. The idea was that he would keep the peace and obviate the danger of early and violent death, but early and violent death is exactly what I and my comrades are now facing, in his service. So it seems that our best bet would be to walk away from the battlefield, leaving the monarch incompetent to do that very thing for which the monarchy was created. Its very function, in other words, requires that there be subjects whose loyalty to it is not just a matter of their belief that it fulfills that function.¹⁷

Craig's example helps to bring out something that the case of Jim does not: this kind of clash between first-order dispositions and higher-order justifications is not really the stuff of fanciful philosophical thought experiments. Rather, it grows out of the conditions of human life itself.

After all, we are beings who must form long-term dispositions to navigate our physical and social worlds given limited cognitive resources. Moreover, collective action is premised on the formation of such dispositions; without them, we could not really *rely* on one another in any deep sense. Craig's subjects are loyal to their King, and the function of this disposition is to override the search for higher-order justifications, precisely because those higher-order justifications can easily erode their all-important instinctive commitment to the shared project. For these (and other) reasons, the idea that we could eliminate this structural feature of human society is almost certainly a fantasy.

But this entails that human ethical practices and institutions can never be fully transparent to those who participate in them since reflective inquiry into their value has the tendency to undermine or destroy them. Or, to put it another way, ethics has no chance of being everything that it seems. This is neither a deliverance of purely metaphysical inquiry nor a revelation produced by looking at our lives from a God's-eye perspective. It is, rather, a form of practical alienation that is faced by any reflective agent in a recognizably human social world.

A case study

At this point, we should address a potentially troublesome question. Surely, if this undermining effect is to be a real concern for us, we should be able to locate it in the world, that is, we should be able to find instances in which an agent taking the outside perspective has seriously threatened to undermine his or her deepest commitments. I take this demand very

seriously: the mere spinning-out of a conceptual and moral-psychological theory which entails that there is a real practical problem facing human beings is never sufficient to show that there is a real practical problem facing human beings.

Fortunately, a recent pair of utilitarian writers has (unwittingly) provided us with a perfect case study. I refer here to Peter Singer and Katarzyna de Lazari-Radek's recent admission—in a book defending Sidgwick's "esoteric" utilitarianism—that the theory itself may well deliver the verdict that people shouldn't write books advocating belief in esoteric utilitarianism. "Arguably," they note in a chapter defending esoteric theory, "we should not even have written this chapter."²⁰ It is not hard to see why, since esoteric utilitarianism is only meant to be believed in secret by a group of elites, and the best way to keep the secret is hardly to have a book defending the theory published by the largest academic press in the English-speaking world. Here, Singer and de Lazari-Radek are confronted with a situational conflict between their personal commitment to an identity-defining project—academic writing—and a certain utilitarian version of the "outside" perspective on that project. Since we are reading these passages, we already know which side won the battle. To justify what looks like a violation of their own professed principles, they write:

[I]n a book on Sidgwick, to fail to discuss the topic of esoteric morality would be to leave the impression that on this issue Sidgwick's stance—and therefore utilitarianism in general—is indefensible. That impression could also have bad consequences.²¹

Williams, I believe, would have been troubled by what is almost certainly a failure of truthfulness here. Singer and de Lazari-Radek seem unwilling to simply admit that they love academic writing, that philosophy is something to which they have devoted their lives, and that the dissemination of true ethical ideas is enormously important to them, consequences be damned. But this is not what they say. Instead, they choose to describe their decision to publish their theory as the outcome of utilitarian reasoning. Moreover, the reasoning itself is hard to process: an esoteric utilitarian, by definition, should *want* most of the population to think that utilitarianism is indefensible. By the lights of the theory, that is a *good* consequence, not a bad one. Because the utilitarian reasoning itself is inadequate, we should conclude that Singer and de Lazari-Radek's motivating reasons were grounded in love and commitment, and not on their acceptance of any ethical theory.

Yet, because the activity they love is so hard to justify by the lights of the particular version of "outside" reflection they have adopted, we should also conclude that wholehearted prioritization of that perspective could very well destroy their project. These are the real-world dangers of the outside perspective, dangers which are only avoided by simply switching off that perspective entirely. Moreover, these effects are not confined to utilitarian philosophers: an analogous practical problem faces any moral philosopher whose preferred conception of right action does not harmonize well with the extraordinary amount of time, energy, and public resources he or she expends in the tortuous search for the correct theory of right action.

This ends my discussion of the skepticism Williams advocated in *ELP*. As the case study helps to show, the skepticism amounts to the claim that a recognizably human agent cannot have a *life* at all while giving deliberative priority to a higher-order or "outside" vindication of that life. As I will now argue, however, Williams' deep skepticism may be his own undoing. If ethics cannot be everything that it seems, then Williams' well-known and influential attacks on moral theory seem to lose a great deal of force.

The inevitability of inauthenticity

In *Utilitarianism:* For and Against, Williams claimed to have identified a way in which utilitarianism alienates agents from their projects. Acceptance of the theory, he claimed, threatens to eradicate the very dispositions around which ethical life is built.²² A very similar thought is implicit in the famous "one thought too many" problem for Kantian ethics, where it is suggested that the search for a higher-order justification for saving one's spouse will undermine one's commitment to the relationship.²³ These arguments have achieved canonical status, and Utilitarians and Kantians have spent a great deal of energy responding to them.

But now, in *ELP*, we have something much more ambitious than a particular argument against a particular ethical theory. What we have is a *global* claim that no reflective inquiry can make ethical thought and practice fully transparent without producing practical alienation. What began its life as a specific critique of J.J.C. Smart's view has become a universal solvent which will destroy any attempt to make ethics stand up to reflection. Since Williams insists that "there is no route back" from the reflective life, he now faces a very serious problem: in what sense is it a *criticism* of an ethical theory that its acceptance produces a scenario which we already inhabit, one from which there is virtually no escape?

From Williams' writings on utilitarianism and on Kantian ethics, one certainly gets the sense that he believed in some alternative, non-alienated state of affairs which was waiting for those of us who shared his insights about the ways in which such theories could distort our lives. In "Persons, Character and Morality," alienation was traced to the "impartial moral consciousness" which feels "unease" at the possibility of conflict with the immediate demands of friendship and love. And in *Utilitarianism: For and Against* alienation was said to be the result of our adopting a "purely utilitarian point of view." Kantian and utilitarian theories are models for how "outside" reflection should work, and Williams claimed that each

results in practical alienation. Yet, in *ELP*, we find him claiming that *any* "outside" reflection must produce that same alienation.

This is a very serious problem, one which threatens to undermine many interesting arguments in *ELP* itself. Consider, for example, Williams' criticism of R.M. Hare's theory. Hare famously recommends two levels of moral thinking, intuitive and critical. Most of the time, Hare permits us to act instinctively on our ethical dispositions. However, some of our time must be spent in what he called the "cool hour," within which we suspend our commitment to those dispositions and critically inquire into their utilitarian value. The passage from *ELP* in which Williams criticizes this view is very revealing:

Is there anywhere in the mind or in society that a theory of this kind can be coherently or acceptably located? The theory finds a value for these dispositions, but it is still an instrumental value... This is what those dispositions look like when seen from outside, from the point of view of the utilitarian consciousness. But it is not what they seem from the inside. Indeed, the utilitarian argument implies that they should not seem like that from the inside. The dispositions help to form the character of an agent who has them, and they will do the job the theory has given them only if the agent does not see his character purely instrumentally, but sees the world from the point of view of that character.²⁵

Notice that this is precisely the same "gap" between inside and outside perspectives that must, according to Williams, destroy any attempt to make ethics stand up to reflection. According to this (putative) criticism, Hare's theory asks us to be like the subjects in Craig's example, where moments of reflection on the true function of our dispositions coexist, in an unstable manner, with our wholehearted commitments. This sounds like a nasty state to be in. But, again, if Williams is right, and no reflective inquiry can make ethical life transparent and still preserve it, then it is the state that any reflective ethical agent must be in if she is to have such a life at all. Williams asks, rhetorically, where such a theory could be "coherently or acceptably located," but the answer to this question is now plain: Hare's theory can be located precisely in the alienated human consciousness, which is doomed to vacillate between conflicting justifications. ²⁶

The point, I hope, has been made clear. One of Williams' most enduring ideas is that certain ethical theories threaten to produce a kind of practical alienation in us. Yet, if there is no alternative to this state, then this is not a criticism of such theories. In fact, it turns out that theories such as Hare's are *realistic* in a highly desirable way: they correctly describe the structure of ordinary ethical experience. What can Williams say in response to this criticism? In the following two sections, I'll canvass two possible lines of response, arguing that neither solves this problem.

The tragedy of truth and truthfulness

Here, those familiar with Williams' work will note that he suggested that at least some of our commitments could be made transparent to reflection in the right way. In *ELP*, he wrote:

While ethical thought will never entirely appear as what it is, and can never fully manifest the fact that it rests in human dispositions, this will present greater obstacles to reflection in some conditions of ethical thought than in others. One thing that will make a difference is the extent to which ethical life can still rely on what I have called thick ethical concepts. They are indeed open to being unseated by reflection, but *to the extent that they survive it*, a practice that uses them is more stable in face of the general, structural reflections about the truth of ethical judgments than a practice that does not use them.²⁷

Indeed, TT is precisely an attempt to show that the thick concept "truthfulness" can survive reflection. Indeed, the book may be nothing less than an attempt to partially rescue us from the sort of deep and pervasive practical alienation I've described here. However, as I now hope to show, Williams' attempt is ultimately tragic in a special sense. After a philosophical lifetime spent casting scorn on a mode of reasoning which alienates us from our projects, Williams finds himself unable to avoid that very mode of reasoning. He faces a problem: how to underwrite or reflectively vindicate our dispositions to truthfulness in a manner that is distinct from the sort of vindication provided by utilitarian such as Hare. I do not think that he solves this problem, at least, not here.

Let us remind ourselves that this mode of reasoning essentially involves instrumentalizing our dispositions, portraying them, in Williams own apt phrase, as "devices for generating states of affairs." A sophisticated consequentialist like Hare will recommend that we generally act on our most familiar dispositions, trusting that they are appropriately responsive to the ethical facts. Yet, they recommend this only because doing so is conducive to the general good, and not because the beliefs implicit in the dispositions themselves are true. We are thus invited to view our own dispositions merely as parts of the causal nexus, and as we have seen, Williams believed that this "outside" perspective produces practical alienation or inauthenticity.

Unfortunately, it is very difficult to see how Williams avoids instrumentalizing truthfulness in precisely the same way. He claims to be doing more, but at crucial moments in the book, we are left with nothing but very general claims about how our dispositions to be sincere and accurate support, in a basically causal sense, other states of affairs that we value.

In the first half of TT, Williams spends some time showing that a fictional genealogy of truthfulness can deliver the conclusion that the virtues of truth are of immense instrumental value for human beings. Of course,

for the reasons just outlined, he recognized that this was not enough, and he went on to suggest that given a suitably sophisticated conception of intrinsic value, the real history of truthfulness could illustrate its intrinsic value to us. Briefly, he wanted us to see an intrinsic value as one which could be coherently related to other things we value, such as scientific activity, non-tyrannical political systems, and even authenticity itself. This relation, again, was supposed to be distinct from the means-ends relation characteristic of instrumental value. By illustrating the social history of these values, he aimed to show that truthfulness is deeply related to those values. He argued that the development of such things as science, democracy, and authenticity was intimately bound up with the increased valuation of truthfulness. This "filling in" of the earlier genealogy with real history was supposed to enable Williams to pull off the requisite trick, to provide an evaluative foundation for truthfulness which could be wholeheartedly accepted by someone with the dispositions itself. ²⁹

The difficulty may be put quite simply: this argument is a paradigmatic example of the "outside" perspective at work, the very perspective which, according to Williams himself, cannot directly vindicate any first-order commitment. The vindication is as indirect as any sophisticated consequentialist account. It doesn't matter if we grant Williams to right to call truthfulness "intrinsically" valuable; nor does it matter that truthfulness is a thick ethical concept. These issues, unfortunately, are orthogonal to the skeptical position I'm considering in this paper. What matters is the structure of the vindication itself. It does not show that it is simply good to be truthful. Rather, it is a higher-order, reflective story which takes very general social goods and relates them to our dispositions to be truthful. It is indeed a vindicatory story, but as the quotation at the outset of this paper makes clear, that doesn't address the problem at all, since Williams' skepticism is supposed to have force even in the presence of a successful vindication of ethics. Recall that even *if* some bit of our ethical thought is vindicated,

the consequences of that could lie only in justifying a disposition to accept certain ethical statements, rather than in showing, directly, the truth of those statements: but this is not how it would naturally appear to those who accepted them.

For my own part, I simply cannot see how Williams circa 2003 avoids the very trap that he himself had constructed for other theorists, circa 1985. This is why I say that TT is a tragic book: in it, Williams becomes ensnared by a certain very subtle trap, one which had been designed and set by a younger version of himself.

That said, it might be suggested that Williams has a different response to this problem, one grounded in his (apparent) commitment to a certain kind of subjectivism about justification. It is to this second strategy that I now turn.

Subjectivism and alienation

In "Moral Luck," Williams described a fictionalized Gauguin who abandons his family to pursue an artistic project. Williams' stated aim in the paper was to "explore and uphold the claim that in such a situation the only thing that will justify [Gauguin's] choice will be success itself." Yet, he immediately qualified this claim:

One should be warned already, however, that, even if Gauguin can be ultimately justified, that need not provide him with any way of justifying himself to others... Thus he may have no way of bringing it about that those who suffer from his decision will have no justified ground of reproach. Even if he succeeds, he will not acquire a right that they accept what he has to say.³⁰

And later, he writes that the family's complaints "are, indeed, justified," even in the case where Gauguin's project succeeds. How are we to make sense of this?

In spite of his refusal to explicitly endorse any such position, one standard interpretive route is to read Williams as a kind of subjectivist about justification. Sharon Street and Derek Parfit, for example, have recently interpreted Williams along these lines.³² The position might look something like this:

Subjectivism: X truly judges that action A is justified iff X has an evaluative framework according to which it is all-things-considered best to perform A, given complete factual knowledge of the situation.³³

But subjectivism, it must be stressed, is a theory about how the "outside" perspective ought to conduct its business. In other words, it is a perfectly general theory, a formula which enables us to determine the normative status of any action whatsoever. Here, I must tangentially mention that this feature of subjectivism explains Williams' refusal to endorse it: subjectivism is a *theory*, and as such it would be odd for this most trenchant of antitheorists to embrace it. Furthermore, contrary to numerous conflations that crop up over and over again in the literature, Williams' internal reasons thesis does not entail subjectivism about reasons.³⁴

Nonetheless, let us grant, for the sake of argument, that he could have held this position. If he did, it would be consistent for him to say, as he seems to say, that Gauguin can rightly feel justified in abandoning his family and that his family need not accept or acknowledge his justification. Each is correct to think as they do, given their respective evaluative frameworks. Moreover, it might be thought that this sort of position can save Williams from the objection I am raising in this paper. The objection, recall, is that Williams has painted himself into something of a skeptical corner.

He criticized various moral theories for producing a certain kind of practical alienation, but he subsequently developed an argument according to which that sort of alienation is inevitable for creatures like us. This, again, is because of a certain mismatch between the deliverances of the "outside" perspective and the natural sense of justification that comes along with taking the "inside" perspective on our own dispositions.

The suggestion under consideration therefore amounts to the following: by embracing a theory of justification which defines correct action in terms of an agent's evaluative scheme, it will never be possible for the outside and inside perspectives to conflict. Since the beliefs implicit in the inside perspective arise directly from the very psychological elements which form the basis of a subjectivist theory of justification, it will (we suppose) always be the case that an agent is justified in following through on their deepest commitments. Jim will feel as though he can't shoot the villager, and he will be justified in refusing to do so. Singer and de Lazari-Radek will feel as though they *must* publish their book, and they will be justified in doing so. The inside and outside perspectives will always harmonize, and as such, subjectivism is a one-shot antidote to practical alienation.

Yet, a further problem looms, one that becomes salient when we take Williams at his word about Gauguin. For *what* is it that Gauguin believes (from the inside) when he is compelled to pursue his art? Certainly not that he desires to be a successful artist and that this desire provides him with all the justification he needs. Rather, the putative justificatory work will be done—as Williams explicitly says it will—by the prospect of *success*. That possible-future state of affairs itself will seem reason-providing, and not his present desire for it. And he is no outlier, here: rarely do agents wholeheartedly pursue patterns of action while thinking that are justified solely by the presence of their desires, or by the fact that they happen to make certain evaluative judgments. Using the example of generosity, Williams himself acknowledged this near-truism:

The characteristic and basic expression of a moral disposition in deliberation is not a premiss which refers to that disposition—it is not the basic characteristic of a generous man's deliberations that they use the premiss 'I am a generous man'... Though the generous man is partly characterised by what goes into his deliberations, it is not that what goes into them are reflections on his generosity.³⁵

Thus, even on a subjectivist theory of justification, the outside perspective virtually guarantees practical alienation, since the theoretical vindication of our dispositions will very often be *completely* out of step with how things "naturally appear" to us when we inhabit those same dispositions. Instead of directing me toward the value that appears to inhere in various persons, objects, or states of affairs, subjectivism redirects agents to their own psychology. This is a paradigm instance of indirect justification, and it is no

less jarring than the utilitarian story, which redirects me toward the general good. If utilitarianism is to be rejected for undermining our integrity, then we must reject subjectivism on precisely the same grounds, since whole-hearted acceptance of subjectivism would rob us of the outwardness or world-directedness that Williams himself argued was essential to integrity.

Having explored what I take to be the two most promising escape routes, and having shown that they do not permit escape, I conclude that the skepticism articulated in *ELP* weakens Williams' highly influential criticisms of particular ethical theories. This is because his negative appraisal of such theories is only justified if they lead to an avoidable state of affairs. However, as Williams' own attempts in *TT* show, broadly naturalistic reflection on the value of our dispositions ends up instrumentalizing those dispositions, and it therefore threatens to alienate agents from their own projects. And as our discussion of normative subjectivism has shown, even a theory which *defines* practical justification in terms of our own commitments will certainly fail to capture the way that ethical life seems to us, from the inside.

Williams ought to have insisted that the only antidote to this inevitable practical alienation is itself practical. That is, all we can do, given his assumptions, is re-inhabit the inside perspective and simply live out our dispositions, refusing to see them as justified only by the lights of a higher-order theory. However, other theorists might wish to take another route and attack his skepticism at its roots. Consequentialists, in particular, may want to question entirely the normative authority of the "inside" perspective on ethics to deny that practical alienation is of any real significance. I'll conclude by arguing that this move is not feasible, since there is an important sense in which the inside perspective is logically prior to the outside one, even for a consequentialist.

The priority of the "inside" perspective

Some ethical theories self-consciously start with the inside perspective. For example, one of Kant's most influential arguments begins with the claim that each agent is committed to the value of his or her own agency, in virtue of having goals at all. He then moves to the claim that we must value any instantiation of human agency as such.³⁶ The move is much-disputed, but its starting point is recognizably the agent's own private perspective on his or her values.

Consequentialists ordinarily avoid giving any special priority to the particular standpoint of any agent. This, of course, is meant to be a strength of this type of theory: by prioritizing the general good, such a theory can avoid giving any pride of place to prejudice, self-indulgence, or subjective arbitrariness. However, as Williams himself pointed out, a *purely* impersonal standpoint is of little help to a consequentialist.

"The good of any one individual," Sidgwick famously wrote, "is of no more importance, from the point of view of the Universe, than the good

of any other."³⁷ This utilitarian's claim, we should notice, is consistent with the following proposition: no individual's good is of any importance. The universe, it seems, has to decide that what is good for agents is good *as such*, and mere impartiality cannot deliver this substantive conclusion.

Here, the consequentialist must, in the current jargon, deploy an *intuition* about which properties or states of affairs have intrinsic value. This much is admitted, for example, by hardline consequentialist Joshua Greene, who concedes that even consequentialism has to be based on an "affectively based evaluative premise." Perhaps he will suggest that the intuition that suffering is intrinsically bad is self-evident. He will invite us to reflect upon cases, real or imaginary, to experience this self-evidence for ourselves. He might easily succeed in this, securing agreement among all careful, imaginative thinkers. We might even be moved toward an acceptance of some form of utilitarianism. Yet, what is important about this kind of argumentative move is that it straightforwardly invokes the "inside" perspective on values. It follows that no consequentialist of this variety can question the normative authority of the inside perspective without undermining his or her own theory.

This may not seem obvious, but it is in fact very straightforward. The outside perspective, as we have seen, involves justifying the value of some object or disposition by *quarantining* its importance, by suspending its weight in practical deliberation and trying to say why it has value by the lights of some distinct set of values. However, by considering certain cases and making pro-utilitarian judgments about them, we are not suspending our commitment to the badness of suffering in order to say why suffering is bad. Rather, the argument has a much more direct form: we simply allow our disposition to judge that suffering is bad to operate normally.

The consequentialist might insist that the badness of suffering is the deliverance of reason and not emotion or sentiment. This, it will be said, gives it a special authority. This maneuver is typical of intuitionist-consequentialists, who follow Sidgwick in describing certain substantive evaluative claims as rational intuitions, or "intuitive propositions of real clearness and certainty." However, this point is entirely orthogonal to the question I am considering. At no point have I suggested that either the inside or outside perspectives is "rational" or "emotional," nor have I denied that our firmest convictions are accompanied by a certain sense of clarity and certainty. What I am suggesting is that this certainty cannot, logically, be delivered by the outside perspective since it is that very certainty which is *suspended* by the outside perspective. Moreover, the distinction itself is exclusive and exhaustive since either one suspends an evaluative commitment in seeking to justify it or one does not. It follows that the fundamental evaluative judgment(s) lying at the base of any consequentialist theory must be the product of the inside perspective, and that it is disastrous for the consequentialist to call that perspective into question.

The argument over consequentialism can proceed from here, and the contours of that debate are by now familiar. However, it should now be clear that no system of ethical thought, consequentialist or otherwise, can avoid a confrontation with Williams' particular brand of skepticism, which is often passed over in favor of more dramatic, metaphysical skepticisms. But Williams' problem arises from a few very simple, intuitive facts. We can take up one of two conflicting perspectives on our values, and neither perspective is discardable, even in principle.

What do we do with these apparent facts? What would *non-alienating* reflection on our values look like, and would we even want such a thing? These are some of the more central questions bequeathed to us by *ELP*, and they remain as pressing today as they did three decades ago, even if recent writings in ethics show little interest in them. I sincerely hope that renewed attention to this book will allow us to see the skeptical problem for what it is, and that moral philosophers will once again think deeply about the ways in which their theoretical activity can come into conflict with the exigencies of human life.

Notes

- 1 Bernard Williams, Ethics and the Limits of Philosophy (Harvard University Press, 1985).
- 2 There is no mention of the remark in reviews by Sam Scheffler, Susan Wolf, or Thomas Nagel. Thomas Nagel, "Ethics and the Limits of Philosophy by Bernard Williams," *Journal of Philosophy* 83, no. 6 (1986): 351–60; Susan Wolf, "The Deflation of Moral Philosophy: Ethics and the Limits of Philosophy. Bernard Williams," *Ethics* 97, no. 4 (1987): 821–33; Samuel Scheffler, "Morality through Thick and Thin a Critical Notice of Ethics and the Limits of Philosophy," *Philosophical Review* 96, no. 3 (1987): 411–34.) One exception here is Blackburn, S. (1986). Making Ends Meet." Philosophical Books 27 (4): 193-203.
- 3 John L. Mackie, Ethics: Inventing Right and Wrong (Penguin, 1977).
- 4 Williams, Ethics and the Limits of Philosophy, 222.
- 5 Thomas Nagel, The View From Nowhere (Oxford University Press, 1986), vol. 37.
- 6 One recent writer, Paul Sagar, has interpreted the so-called "gap" between inside and outside perspectives as, roughly, that infamous chasm between the amoralist's actual reasons and the reasons we wish that he had. But, this interpretation cannot be right since it fails to make sense of why the gap is supposed to be a problem *for us* irrespective of our dealings with Sensible Knaves, of why it is supposed to produce a form of practical alienation even for those of us who *do* live recognizably ethical lives. The gap, I will argue, runs much deeper than this. See Paul Sagar, "Minding the Gap: Bernard Williams and David Hume on Living an Ethical Life," *Journal of Moral Philosophy* 11, no. 5 (2014): 615–38.
- 7 From *Truth and Truthfulness:* "one way in which others can come to make sense of the action is by thinking themselves into the position of the agent, while taking on for the purpose of the exercise, so far as they can, his outlook and preconceptions... His identification with the outlook of the agent is temporary and, as it were, feigned... In current jargon, the exercise is conducted 'off-line'" (TT 237). See Alvin Goldman, *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (Oxford University Press, 2006), vol. 144.

- 8 Williams would not, of course, be the first philosopher to develop a broadly *social* model of impersonal reflection in ethics. The notion was first defended in detail by Adam Smith in *A Theory of Moral Sentiments*.
- 9 "There is simply no conceivable exercise that consists in stepping completely outside myself and from that point of view evaluating *in toto* the dispositions, projects, and affections that constitute the substance of my own life." Bernard Williams, "The Point of View of the Universe: Sidgwick and the Ambitions of Ethics," in *The Sense of the Past: Essays in the History of Philosophy* (Princeton University Press, 2009), 277–96.
- 10 This set may even include the importance of loyalty in general, from which one still cannot derive the truth of "I ought to be loyal to this particular person."
- 11 Bernard Williams, *Shame and Necessity* (University of California Press, 1992), 75–85.
- 12 John Jamieson Carswell Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge University Press, 1973).
- 13 Arguing against the utilitarian's invocation of "remoter effects" was Williams' way of securing this conclusion.
- 14 Bernard Williams, *Truth and Truthfulness: An Essay in Genealogy* (Princeton: Princeton University Press, 2002).
- 15 Bernard Williams, "Practical Necessity," in *The Philosophical Frontiers* of Christian Theology: Essays Presented to D.M. Mackinnon, eds. Donald MacKenzie MacKinnon, Brian Hebblethwaite, and Stewart R. Sutherland (Cambridge University Press, 1982). For an illuminating discussion, see Robert J. Gay, "Bernard Williams on Practical Necessity," *Mind* 98, no. 392 (1989): 551–69.
- 16 Bernard Williams, Making Sense of Humanity (Cambridge University Press, 1995), 169.
- 17 Edward Craig, "Genealogies and the State of Nature," in *Bernard Williams*, ed. Alan Thomas (Cambridge University Press, 2007), 181–200.
- 18 See Bratman 1999. For a reading of Williams that brings him very close to Bratman in this way, see Markovits, Daniel, "The Architecture of Integrity", 129–32.
- 19 Michael Bratman, Structures of Agency: Essays (Oxford University Press, 2007); Faces of Intention: Selected Essays on Intention and Agency (Cambridge University Press, 1999).
- 20 Katarzyna de Lazari-Radek and Peter Singer, *The Point of View of the Universe: Sidgwick and Contemporary Ethics* (Oxford University Press, 2014).
- 21 Ibid.
- 22 Smart and Williams, *Utilitarianism: For and Against.*
- 23 Bernard Williams, "Persons, Character, and Morality," in *Moral Luck*, ed. James Rachels (Cambridge University Press, 1981).
- 24 Ibid., 17.
- 25 Williams, Ethics and the Limits of Philosophy, 108.
- 26 In response to Williams' complaint, Hare writes:

I do my own moral thinking in the way described in this book... doing my best to employ critical and intuitive thinking as appropriate. In difficult situations one's intuitions, reinforced by the dispositions that go with them, pull one in different directions, and critical thinking, perhaps, in another.

Richard Mervyn Hare, Moral Thinking: Its Levels, Method, and Point (Oxford University Press, 1981), vol. 93, 52.

- 27 Williams, Ethics and the Limits of Philosophy, 261.
- 28 The conception of intrinsic value is briefly laid out and defended in Bernard Williams, "Plato's Construction of Intrinsic Goodness," in *The Sense of the*

Past: Essays in the History of Philosophy, ed. Adrian W. Moore (Princeton University Press, 2009), 118–37. A similar account of intrinsic value can be found in Kagan (1998), "Rethinking Intrinsic Value." Kagan argues, to my mind convincingly, that the fact that an object has value in virtue of its relational properties does not entail that it is only valuable as a means to promoting its relata. The pen that Abraham Lincoln used to sign the emancipation proclamation possesses intrinsic value, and the explanation of this value makes reference to the fact that it was used for a certain purpose by a certain man, Yet, that explanation in no way implies that the pen is only valuable because it can be used to sign important documents since its destruction and replacement by a similar pen would result in a diminishment of value in the world. A similar line of argument can be found in Korsgaard (1983). See Shelly Kagan, "Rethinking Intrinsic Value," Journal of Ethics 2, no. 4 (1998): 277–97; Christine M. Korsgaard, "Two Distinctions in Goodness," Philosophical Review 92, no. 2 (1983): 169–95.

- 29 TT, Chapters 5-8.
- 30 Bernard Williams, Moral Luck: Philosophical Papers, 1973–1980 (Cambridge University Press, 1981), 23–24.
- 31 Williams, Moral Luck: Philosophical Papers, 1973-1980, 37.
- 32 Derek Parfit, On What Matters (Oxford University Press, 2011), vol. II, 270; Sharon Street, "Objectivity and Truth: You'd Better Rethink It." (unpublished ms) Available at https://files.nyu.edu/ss194/public/sharonstreet/Writing.html
- 33 Numerous qualifiers would have to be added here to generate a reasonably complete and defensible model of normative subjectivism. I omit these for simplicity's sake since the point I wish to make in this section does not turn on the presence of any such qualifiers.
- 34 The internal reasons thesis, recall, is the view that a *necessary* condition on an agent having a reason for action is that it be suitably related to the agent's subjective motivational set. Absent a sufficiency condition, internalism cannot entail subjectivism. The conflation of internalism with normative subjectivism is, unfortunately, extremely common in the meta-ethical literature. See Parfit (op. cit); Chris Heathwood, "Desire-Based Theories of Reasons, Pleasure, and Welfare," Oxford Studies in Metaethics 6 (2011): 82; Ralph Wedgwood, "Instrumental Rationality," Oxford Studies in Metaethics 6 (2011): 292; William Fitzpatrick, "Robust Ethical Realism, Non-Naturalism, and Normativity," Oxford Studies in Metaethics 3 (2008): 165. For an admirably clear illustration of the distinction between internalism and subjectivism, see Kate Manne, "Internalism about Reasons: Sad but True?" Philosophical Studies 167, no. 1 (2014), 106.
- 35 Williams, Moral Luck: Philosophical Papers, 1973–1980, 48–49.
- 36 Immanuel Kant, *Groundwork for the Metaphysics of Morals* (Oxford University Press, 1785) AK. 4:429.
- 37 Henry Sidgwick, Methods of Ethics (Kaplan Publishing, 1901), 253.
- 38 Joshua D. Greene. "Beyond Point-and-Shoot Morality: Why Cognitive (Neuro) Science Matters for Ethics." *Ethics* 124, no. 4 (2014): 695–726.

Works cited

Bratman, Michael. Faces of Intention: Selected Essays on Intention and Agency. Cambridge University Press, 1999.

———. Structures of Agency: Essays. Oxford University Press, 2007.

Blackburn, Simon. (1986). "Making Ends Meet." *Philosophical Books* 27(4): 193-203.

- Craig, Edward. "Genealogies and the State of Nature." In *Bernard Williams*, edited by Alan Thomas, 181–200. Cambridge University Press, 2007.
- Fitzpatrick, William. "Robust Ethical Realism, Non-Naturalism, and Normativity." Oxford Studies in Metaethics 3 (2008): 159–205.
- Gay, Robert J. "Bernard Williams on Practical Necessity." *Mind* 98, no. 392 (1989): 551–69.
- Goldman, Alvin. Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading. Vol. 144. Oxford University Press, 2006.
- Greene, Joshua D. "Beyond Point-and-Shoot Morality: Why Cognitive (Neuro) Science Matters for Ethics." *Ethics* 124, no. 4 (2014): 695–726.
- Hare, Richard Mervyn. Moral Thinking: Its Levels, Method, and Point. Vol. 93. Oxford University Press, 1981.
- Heathwood, Chris. "Desire-Based Theories of Reasons, Pleasure, and Welfare." Oxford Studies in Metaethics 6 (2011): 79–106.
- Kagan, Shelly. "Rethinking Intrinsic Value." *Journal of Ethics* 2, no. 4 (1998): 277–97. Kant, Immanuel. *Groundwork for the Metaphysics of Morals*. Oxford University Press, 1785.
- Korsgaard, Christine M. "Two Distinctions in Goodness." *Philosophical Review* 92, no. 2 (1983): 169–95.
- Lazari-Radek, Katarzyna de, and Peter Singer. *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. Oxford University Press, 2014.
- Mackie, John L. Ethics: Inventing Right and Wrong. Penguin, 1977.
- Manne, Kate. "Internalism about Reasons: Sad but True?" *Philosophical Studies* 167, no. 1 (2014): 89–117.
- Nagel, Thomas. "Ethics and the Limits of Philosophy by Bernard Williams." *Journal of Philosophy* 83, no. 6 (1986): 351–60.
- ——. The View From Nowhere. Vol. 37. Oxford University Press, 1986.
- Parfit, Derek. On What Matters. Oxford University Press, 2011.
- Sagar, Paul, "Minding the Gap: Bernard Williams and David Hume on Living an Ethical Life," *Journal of Moral Philosophy* 11, no. 5 (2014): 615–38. Scheffler, Samuel. "Morality through Thick and Thin a Critical Notice of Ethics and the Limits of Philosophy." *Philosophical Review* 96, no. 3 (1987): 411–34.
- Sidgwick, Henry. Methods of Ethics. Kaplan Publishing, 1901.
- Smart, John Jamieson Carswell, and Bernard Williams. *Utilitarianism: For and Against*. Cambridge University Press, 1973.
- Wedgwood, Ralph. "Instrumental Rationality." Oxford Studies in Metaethics 6 (2011): 280-309.
- Williams, Bernard. Making Sense of Humanity. Cambridge University Press, 1995.
- ——. "Persons, Character, and Morality." In *Moral Luck*, edited by James Rachels. Cambridge University Press, 1981, pp. 1–19
- ——. "Plato's Construction of Intrinsic Goodness." In *The Sense of the Past: Essays in the History of Philosophy*, edited by Adrian W. Moore, 118–37. Princeton University Press, 2009.
- ——. "Practical Necessity." In *Moral Luck*, edited by James Rachels. Cambridge University Press, 1981, pp. 124–31.
- ———. *Shame and Necessity*. University of California Press, 1992.
- -----. "The Point of View of the Universe: Sidgwick and the Ambitions of Ethics." In *The Sense of the Past: Essays in the History of Philosophy*, edited by Adrian Moore, 277–96. Princeton University Press, 2009.

208 Nicholas Smyth

- Ethics and the Limits of Philosophy. Harvard University Press, 1985.
- Wolf, Susan. "The Deflation of Moral Philosophy: Ethics and the Limits of Philosophy. Bernard Williams." *Ethics* 97, no. 4 (1987): 821–33.