

Internal Reasons and the Integrity of Blame*

In “A Critique of Utilitarianism” (found in *Utilitarianism, For and Against*), Bernard Williams argues against utilitarianism, claiming it cannot accommodate an agent’s integrity. This argument is underwritten by Williams’ views on reasons, found in his article “Internal and External Reasons.” In another article, “Internal Reasons and the Obscurity of Blame,” Williams attempts to reconcile his views on reasons with the practice of blame. (Both “Internal and External Reasons” and “Internal Reasons and the Integrity of Blame” have been recently reprinted in *Making Sense of Humanity and Other Philosophical Papers*, Cambridge University Press, 1995.) That attempt is not only unsuccessful, but actually falls prey to his own argument against utilitarianism. In examining this inconsistency, we will discover how accommodating blame limits (but does not eliminate) our ability to criticize an ethical theory for neglecting the particularities of the agent.

1. The Two Earlier Positions

1.1

In “A Critique of Utilitarianism” Bernard Williams argues that utilitarianism “cannot hope to make sense, at any serious level, of integrity... [because] it can make only the most superficial sense of human desire and action at all” (1973, p.82). His point is not merely that utilitarianism, with its emphasis on happiness, does not properly value personal integrity. Williams concedes that even if utilitarianism “makes integrity as a value more or less unintelligible... it would not necessarily follow that we should reject utilitarianism; perhaps... we should just forget about integrity, in favor of such things as a concern for the common good” (1973, p.99). The problem Williams sees is deeper: “the reason why utilitarianism cannot understand integrity is that it cannot coherently describe the relations between a man’s projects and his actions” (1973, p.100):

The point is that [the agent] is identified with his actions as flowing from projects and attitudes which in some cases he takes seriously at the deepest level, as what his life is about... It is absurd to demand of such a man, when the sums come in from the utility network... that he should just step aside from his own project and decision and acknowledge the decision which utilitarian calculation requires. It is to alienate him in a real sense from his actions and the source of his action in his own convictions. It is to make him into a channel between the input of everyone's projects, including his own, and an output of optimific decision; but this is to neglect the extent to which *his* actions and *his* decisions have to be seen as the actions and decisions which flow from the projects and attitudes with which he is most closely identified. It is thus, in the most literal sense, an attack on his integrity. (1973, p.116-117)

Utilitarianism cannot understand integrity as a value because it does not understand the relation between a person's projects and his actions. This lack of understanding leads the utilitarian to demand that the agent do things which the agent, in light of his own projects and convictions, has no *reason* to do.

1.2

Williams has specific views about when an agent can be said to have a reason to act, views which depend upon his understanding of the relationship between a person's projects and actions. In "Internal and External Reasons" Williams presents two interpretations for the statement "A has a reason to ϕ ," labeling them "internal" and "external."¹ A reason statement is internal, roughly, if the reason given is rightly related to some element in the agent's "subjective motivational set"² (which Williams calls his "S"), and external if it bears no such relation. Williams illustrates his distinction with the following example:

When we say a person has reason to take medicine which he needs, although he consistently and persuasively denies any interest in preserving his health, we may well still be speaking in the internal sense, with the thought that really, at some level he *must* want to be well. However, if we become clear that we have no such thought, and persist in saying that the person has this reason, then we must be speaking in another sense, and this is the external sense. (1981, p.106)

Williams argues that in the latter case we are mistaken: there is no reason for the agent to take the medicine.

To establish this, Williams first argues that nothing can be considered a reason unless it is able to motivate the agent (1981, p.106f). That argument will not be considered here.³ For our purposes simply note that he does *not* argue that a consideration counts as a reason only if it *actually* motivates the agent, rather only if it *is able to*. Reasons statements must play both an explanatory and a normative role (1989, p.2). Thus *all* reasons are, simply by virtue of being reasons, at least *potentially* motivational, while only some of them are actually so.

Any reason that *actually* motivates an agent is, of course, traceable to the agent's motivational set, and therefore an *internal* reason (1989, p.5). Williams argues that any reason that is even *potentially* motivational, is also, by virtue of that potentiality, traceable to the agent's S, and therefore internal. Since *all* reasons are either actually or potentially motivational, there are no external reasons.

When arguing that any potentially motivational reason is traceable to the agent's S, Williams relies on two crucial premises: that the agent must come to be motivated by a reason through a process of rational deliberation (1981, p.108f) and that all rational deliberation is responsible to one's S (1981, p.104f).⁴ With these premises in place, his argument is sound: the only reasons which are even *potentially* motivational are those to which we can construct a rational deliberative route from the agent's S, i.e. internal reasons.

The first of these two premises importantly affects our understanding of the internalist position, giving us a hint of the problems that will develop later: when determining whether a consideration that does not actually motivate the agent is nevertheless potentially motivational (and therefore a genuine reason), the internal reasons theorist hypothetically corrects the agent's process of *deliberation*. If the consideration then produces the motivation, it counts as a reason. If not, it does not. *Only* the agent's process of deliberation can be corrected, nothing else. Correcting someone's process of deliberation "involves... at least correcting any errors of fact and

reasoning involved in the agent's view of the matter" (1989, p.2). It can also include excursions of imagination (1981, p.104). But it does *not* allow for a third person prediction of the effects of genuine experience⁵, nor does it allow correcting the agent's moral or prudential views. This restriction is crucial, because

if we were to allowed to adjust the agent's prudential and moral assumptions to some assumed normative standard, then obviously there would be no significant difference between the internalist and the externalist accounts. (1989, p.3)

Thus, the internalist position relies upon a distinction between the notion of a "sound deliberative route" (interpreted as "at least" correct facts and reasoning) and other normative considerations. The reason must be able to motivate the agent where "able to" means "able to if he practiced sound rational deliberation" and not "able to if his experience or character were such and such."

1.3

One of the main targets of Williams' argument is the external reasons theorist who wants to level a charge of *irrationality* against the agent who is unmotivated by a proffered reason. Williams' argument shows that this charge is bluff (1981, p.110-1). To call the unmotivated agent irrational is, on Williams view, simply browbeating. It is to neglect the extent to which the agent's deliberation and decisions flow from the members of his S: the agent is not irrational; he simply has no reason to act and nothing in his S from which to rationally deliberate and reach such a reason.

Now we can see how Williams' internalism underwrites his integrity objection. He said that utilitarianism does not understand the relationship between the agent and his actions. It neglects the extent to which the agent's actions and decisions have to be seen as flowing from the projects and attitudes with which he is most closely identified. Understanding Williams' internalism, we can now see that the projects and attitudes with which the agent is most closely identified provide the dominant members of his S, from which flows both deliberation and action. And yet utilitarianism tells the agent to treat his own projects and attitudes in an impartial way, to view them as if they

belonged to someone else. This is, in Williams' words, absurd. The agent can neither act nor deliberate from a truly impartial standpoint, as 'anyone.' The demand alienates him, in a real sense, from the source of his actions and deliberation. He has "no motivation to deliberate from" (1981, p.109, italics deleted) and no (internal) reason to do the things required.

The utilitarian, like the imagined external reasons theorist, browbeats the agent, ignores the relationship between his projects and his actions, and thus attacks his integrity.⁶ The utilitarian demands, on pains of a charge of *immorality*, that the agent look at himself impartially and act in accord with the utilitarian principle. The real danger arises if this browbeating is effective. If the agent has in his S a disposition to avoid browbeating, or perhaps a disposition to do that which those he trusts have convinced him is moral, then he may be motivated by a bare normative consideration, a bare "ought," at best loosely related to the projects and attitudes with which the agent is most closely identified. He might be compelled by this one disposition to "just step aside" from his own deeper concerns and do that which it requires. He would be acting in a way in which, on balance, he has no reason to act. He would be compromising his integrity.

2. Problems with Blame

Williams' position cannot accommodate much of our blame. Assuming that "A has a reason to ϕ " has normative force only with regard to sound rational deliberation allows us to blame people only for failures of deliberation, but not for more general failures of character. In light of Williams' arguments, much of our blame, like the externalist's charge of irrationality and the utilitarian's charge of immorality, is mere bluff and browbeating. Williams attempts to accommodate blame in "Internal Reasons and the Obscurity of Blame," but does so in a way that neither allows for genuine criticisms of character nor saves blame from being construed as mere browbeating.

2.1

In explaining the problem, Williams first establishes the connection between reasons and blame: "If 'ought to have' is appropriate afterwards [as] blame, then (roughly) 'ought to' was appropriate at the time [as] advice" (1989, p.7). An ascription of blame is the *post hoc* equivalent of a reason statement.⁷ But, if a reason can only be ascribed when it bears the proper deliberative relation to someone's S, then blame, as well, has to be thus restricted.

[I]t may seem a rather obvious fact about blame that someone can be blamed even though his S does not contain anything that would lead to the appropriate motivations: we can blame a man (we may think) for neglecting his wife even though he has no motivations to be concerned about his wife. So if blame is necessarily connected with reasons, it seems to be necessarily connected with external reasons. (1989, p.7, footnote deleted)

The internalist is in trouble: it seems blame can be external, and if reasons parallel blame, then reasons can be external, too.

The problem is not simply that blame is not *tied* to the agent's S. In fact, blame *is* tied to the agent's S. We often blame someone more for neglecting those things to which he is deeply committed, less for those he is not.⁸ We will often blame someone for compromising his integrity ("How could he?! After all those years of commitment, to just drop it all for the money!"). And we will often blame one person for something for which we would not blame another ("...but in *her* case I would have expected better!"). Blame does not treat the agent as "anyone." The problem blame poses for the internalist is not that blame is completely *indifferent* to the agent's S, but that it seems to be *external* to the agent's S – ascribable even when there is no member of the agent's S that could have motivated the agent.

Williams shuffles around this problem in a way that is consistent with internalism, but inconsistent both with the way we actually ascribe blame and with the spirit of his integrity argument. In order for blame to be internal, it needs to address some member of the agent's S. Williams calls our attention to "the ethically important disposition that consists in a desire to be respected by people whom in turn one respects" (1989, p.7).

He believes that most people have this disposition, and that it provides the needed motivation to avoid blameworthy action in what Williams calls a “proleptic mechanism.”⁹ The desire to be respected provides the agent with an internal reason to avoid those things for which he would be blamed in the future, even if he had no other reason to avoid them (1989, p.8). The blame addresses this disposition, and thus is internal blame.

Williams does not ascribe this disposition to *everyone's* S. Blame statements are, in this respect, like “optimistic internal reasons statements;” we launch them hoping that somewhere in the agent is some motivation that might issue in the action we seek (1989, p.6). If someone hasn't even this disposition to avoid blame, then blame *would* be genuinely external. Williams calls him a “hard case,” and claims that

it is precisely people who are regarded as lacking any general disposition to respect the reactions of others that we cease to blame, and regard as hopeless or dangerous characters, rather than thinking that blame is appropriate to them. (1989, p.9)

Williams is consistent with his internalism: there is no such thing as external blame.

2.2

But Williams is inconsistent with the way we actually ascribe blame: we *do* ascribe external blame, and his own example shows that clearly. We might blame a man for not allowing his desires, e.g. ‘to be respected by those he respects,’ to inform his actions (although we might blame him for just the opposite, calling him a spineless hypocrite). We also might blame him for not being nice to his wife. But these are surely two very *different* accusations. Blaming someone for failing to *recognize* some motivation he had and should have acted upon differs greatly from blaming someone for not *having* a motivation he should have. The first is a fault of something like reasoning; the second is a fault of character. Thus we launch external blame statements that are *not* “optimistic” internal blame statements, hoping to reach some motivation. We launch

external blame statements precisely *because* we believe the agent does *not* have the motive that will issue in the action we seek.

Furthermore, we do not hope to find some *other* motive that would issue in the action, *especially* not our blame. Against Williams' assertion about "hard cases," we *blame* people for being unresponsive to blame ("she's *shameless!*"), but we would *also* blame someone for being nice to his wife merely in order to avoid blame. We would call such a man a hypocrite; he is not really being nice to her, he is only concerned with his own image, conforming to our expectations. In blaming him, we do *not* want him to be motivated by our blame, nor are we blaming him for not taking into account the force of our blame; we are blaming him for not taking into account, or not being able to take into account, the force of the reasons for being nice to his wife. Our actual practice of finding fault does not operate on anything like a proleptic mechanism. Williams' attempt to reconcile his internalism with blame fails. Sometimes our blame simply *is* external.

Compare the negligent husband with Olivia, who is committed to non-standard deliberative practices. Olivia interprets all evidence in light of and in support of her belief that most of modern history has been controlled by the CIA. It's all a plot, one big conspiracy. Olivia does not fall into self-contradiction; she merely uses hypothesis which we would describe as silly or *ad hoc*.¹⁰ Olivia is presumably committed to the "respect of her rational peers," but *that* does not back our criticisms of her, nor does it provide her with a proper justification for changing her deliberative practices. When we criticize her silliness, we do so not because it fails to conform to our standards and therefore gains our disrespect, but because, in failing to conform to our standards, we think it misses the truth. In fact, there needs to be some prior judgement, like "it misses the truth," to ground our disrespect in the first place.

Furthermore, we don't consider our criticism a reason for her to adopt our standards of rational practice. At most our criticism provides her with a reason to re-evaluate her

practices to see if they accurately track the truth. If she were to adopt our standards merely out of peer pressure, we would think she was missing their significance and question whether she really understood them at all. We want her to understand the reasons *behind* the practices, not just that they are accepted.

Moral blame works the same way. To blame someone is to say that his response did not appropriately track the circumstance. When we criticize the hard-hearted husband, we do so not because he should have known that failing to conform to our standards would earn our blame (which he presumably wants to avoid). Rather, we criticize him because he either does not understand the force of the reasons or was not affected by the force of the reasons for which we think he should act. If the husband conforms to our expectations out of peer pressure, he is missing the significance of the considerations presented. We would question whether he really understood them at all. We want him to understand the reasons *behind* the practices, not just dryly conform to them.

3. The Integrity Inconsistency

Williams' internalism is inconsistent with the actual practice of blame.¹¹ But it does not necessarily follow that we should reject internalism. Perhaps we should just forget about our usual practice of blame, in favor of Williams' proleptic blame and concern for such things as integrity. Unfortunately, Williams' proleptic blame is itself inconsistent with his concern for integrity.

Consider again Williams' negligent husband. Suppose he is closely identified with his negligent attitude: it is well integrated into his character and his life. We blame him for being so identified, and demand an overhaul of his character, a restructuring of his life. What if we just forgot about external blame, in favor of preserving the man's integrity? We cannot necessarily fall back on proleptic blame because it, too, might attack his integrity.

Proleptic blame protects the integrity of the “hard case;” he alone is exempt from blame. But whether or not one is a hard case depends upon whether or not one is responsive to blame, while whether or not blame attacks one’s integrity depends upon whether or not one is closely identified with the attitude or action for which one is blamed. The negligent husband may be responsive to blame, and thus not a hard case, but nevertheless be closely identified with his negligent attitude, and thus in danger of compromising his integrity. In such a case Williams’ proleptic blame would, just like external blame (and utilitarianism), attack his integrity: the husband is to “just step aside from” these attitudes and do that which our respect requires. Internal, proleptic blame, as much as genuine external blame, “neglects” the extent to which a person’s actions and decisions have to be seen as flowing from the projects and attitudes with which he is *most closely* identified. Williams has here advocated precisely the sort of browbeating for which he criticized the utilitarian and the externalist: he has employed a bare normative consideration in attempt to motivate. Utilitarianism, external blame, *and* proleptic blame all ask the agent to be motivated by considerations of duty or expectation that have no necessary connection to the projects and attitudes with which the agent is most closely identified. Williams is deeply inconsistent. He attempts to salvage his internalism using the same mechanism he criticized with it.

Perhaps more significantly, we have discovered that blame (proleptic or not), like utilitarianism and the external reasons theorist, often asks the agent to compromise his integrity. This is a problem not only for Williams; it must be confronted by anyone who wishes to criticize ethical theories because they are too impersonal and ask the agent to treat himself as “anyone.” Blame is unresponsive to individual particularities in a way very similar to the ethical theories—similar, but not identical. Recall that blame, while able to be external to the agent’s *S*, is nonetheless not independent of it. Blame does not treat the agent as “anyone.” There is still something in the integrity objection: there is room to criticize utilitarianism for being indifferent to the individual,

whereas blame is merely external to the agent's S. But this will be no minor revision to the integrity objection, and will require a position far less subjective than Williams'.

Bibliography

- Dworkin, Gerald (1988). "Autonomy, Science and Morality," in *The Inner Citadel*, John Christman, ed. (Oxford: Oxford University Press), 234-245.
- Hooker, Brad (1989). "Williams' Argument Against External Reasons," *Analysis* 47: 42-44.
- Korsgaard, Christine (1986). "Skepticism about Practical Reason," *The Journal of Philosophy* 83: 5-25.
- McDowell, John (unpublished manuscript). "Might There Be External Reasons?," forthcoming in *World, Mind and Ethics: Essays on the Moral Philosophy of Bernard Williams*, J.E.J. Alther and Ross Harrison, eds.. (Cambridge, Cambridge University Press).
- Millgram, Elijah (in preparation). "Williams Argument Against External Reasons." Princeton University.
- Pincoffs, Edmund (1986). *Quandaries and Virtues*. (Kansas: University of Kansas Press).
- Williams, Bernard (1973a). "A Critique of Utilitarianism," in *Utilitarianism: For and Against*, J.J.C. Smart and Bernard Williams. (Cambridge: Cambridge University Press), 75-150.
- Williams, Bernard (1981). "Internal and External Reasons," in *Moral Luck* (Cambridge: Cambridge University Press).
- Williams, Bernard (1985). *Ethics and the Limits of Philosophy*. (Cambridge, Massachusetts: Harvard University Press).
- Williams, Bernard (1989). "Internal Reasons and the Obscurity of Blame," *Logos* 10: 1-11.
- Williams, Bernard (1995). *Making Sense of Humanity and Other Philosophical Papers*. (Cambridge: Cambridge University Press).

* I am grateful to Elijah Millgram for his advice and encouragement.

¹Williams uses the terms 'internal reasons' and 'external reasons' as a convenience for talking about reasons *statements* (1981, p.101). I will do the same.

²Elements of an agent's subjective motivational set may be termed 'desires,' so long as that term is loosely applied: "[the subjective motivational set] can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying the commitments of the agent." (1981, p.105)

³For discussions of that argument, see Hooker(1989), Korsgaard(1986), McDowell(forthcoming), and Millgram(in preparation).

⁴Williams' premises have been questioned in the discussions mentioned in note 3.

⁵see Millgram (in preparation).

⁶Integrity, in Williams usage, means much more than simple honesty. The projects and attitudes with which one is most closely identified provide the structure not only of one's S, but (therefore) also of one's character and one's life. They need to be consistently and coherently integrated in a way that provides strength of both character and motivation. Compromising one's integrity in Williams' sense is less like tainting one's honesty and closer to the sense in which faulty construction compromises the integrity of a building.

⁷I, like Williams, "am going to assume... that the parallelism between blame and advice does establish a close connection between... blame and the agent's reasons" (1989, p.7).

⁸See Pincoffs (1986), p.22.

⁹"Proleptic" is from "prolepsis... the representation or assumption of a future act or development as being presently existing or accomplished" (from *Webster's Third International Dictionary*).

¹⁰see Dworkin (1988), p.236.

¹¹I believe Williams' problems with blame arise because of the way he has restricted the normative force of reason statements. As mentioned earlier, the internalist position relies upon a distinction between considerations of "fact and reasoning" and moral or prudential concerns. Williams justifies this distinction by saying that we are entitled to assume any rationally deliberative agent has in his S a general desire to be factually and rationally correctly informed. However, we are not similarly entitled to assume that he has in his S any given moral or prudential concerns. Facts and reasonings provide "enough for the notion to be normative" (1989, p.3). Thus the normative force of a reason statement is ultimately grounded in the subjective motivational set of the agent. For problems with this, see Korsgaard (1986).