

The Obscurity of Internal Reasons

Stephen Finlay

University of Southern California

© 2009 Stephen Finlay

<www.philosophersimprint.org/009007/>

SINCE ITS PUBLICATION in 1979, Bernard Williams' "Internal and External Reasons"¹ has been one of the most influential and widely discussed papers in ethics. I suggest here that the paper's argument has nevertheless been universally misunderstood.² On the standard interpretations, his argument — which he subsequently elaborated and defended in further discussions — is perplexingly weak. In the first section I sketch the main variants of the Standard (or, more provocatively, "Supposed") Argument, and catalog the numerous flaws that others have observed in them. The badness of the argument itself might not be a conclusive reason against ascribing it even to a great philosopher, but there are many other exegetical problems for this reading, including that Williams himself seemed blithely aware of and untroubled by many of these flaws.

In the second section I propose a different interpretation of Williams as offering an Alternative (or more provocatively, the "Actual") Argument, which is immune to the seemingly fatal objections to the Standard Argument and better supported by the text and charity.³ On this interpretation, Williams offers just one consistent argument, which unites his central concerns with (i) the "explanatory dimension" of reasons statements, (ii) their conceptual content, and (iii) the connection between reasons and deliberation. While his argument is normally thought to be based on the common claim that reasons must be capable of motivating, I argue that it rather begins from a substantive and interesting analysis of the *concept* of a normative reason: that to believe that *R* is for you a reason for action just is to believe that *R*

1. Henceforth "IER". All references by page number alone are to the 1981 reprint of this paper. Obscurity even surrounds the date of its original publication. Although almost universally cited as from 1980, it actually appeared in 1979. Williams' own source note in his *Moral Luck* (1981) seems to be the origin of the error.
2. Alan Thomas (2002) is a possible exception. Thomas advances an account sharing some of the features of mine, but his sketch is brief and he doesn't argue for his interpretation in detail.
3. Readers uninterested in my painstaking examination of the problems with the standard reading are able to skip straight to the second section for my proposed alternative.

is a certain kind of explanation of action, an explanation of why you would act if you were to deliberate soundly.

1. The Standard (or Supposed) Argument

According to orthodoxy, Williams' argument has approximately the following schematic form:

For all R, A, ϕ ;

- S1. R is a reason for A to ϕ only if A can be motivated to ϕ by believing R ;
- S2. A can be motivated to ϕ by believing R only if there is something in A 's motivational set that could bring A to be motivated to ϕ by believing R ;
- S3. If R is an external reason for A to ϕ , then there is nothing in A 's motivational set that could bring A to be motivated to ϕ by believing R ;
- S4. Therefore there are no external reasons.

On close inspection this argument is puzzling and ambiguous, and raises at least two interpretive questions: (i) What is the intended sense of the *possibility* of A 's being motivated, which links premises S1 and S2? (ii) What exactly does Williams mean by an agent's "motivational set", which links premises S2 and S3? Because S3 is functionally a definition of "external reasons", this question also concerns the precise target of Williams' skepticism.

Each of these issues allows two significantly different readings, which combine to yield four different versions of the Standard Argument. In this section I sketch these readings, critically discuss the four versions of the Standard Argument, and discuss the interpretive problems facing them. Although some interpretations resembling the Standard Argument may elude this net, these problems motivate and direct the search for a more faithful reconstruction of Williams' argument.

(i) Claiming that A can be motivated to ϕ by accepting R is equivalent to claiming that A would under certain conditions be motivated to

ϕ by accepting R . Specifying these conditions identifies the relevant sense of possibility. For the argument to be valid both S1 and S2 must involve the same sense of possibility. The argument's success requires a sense that is weak enough for S1 to be plausible, but strong enough for S2 to be plausible. For example, consider the strong sense of possibility provided by the condition that A has exactly his actual motivational set:

- S1*. R is a reason for A to ϕ only if, under conditions of having his actual motivational set, A would be motivated to ϕ by believing R ;
- S2*. A would be motivated to ϕ by believing R under conditions of having A 's actual motivational set only if there is something in A 's actual motivational set that could bring A to be motivated to ϕ by believing R .

On this reading S2* is weak enough to seem plausible, but even Williams thinks S1* is implausibly strong. He allows, for example, that we can have reasons that do not motivate us only because we have not sufficiently exercised imagination in ways that would lead to new motivations.⁴

In IER, Williams describes the relevant conditions as involving "rationality" ("The internal reasons conception is concerned with the agent's rationality", p. 103). This remains ambiguous, however, because the concept of rationality is vague and contested. To fix terminology I will use 'rationality' only in a very broad sense, as meaning *the dispositions of thought that agents ought ideally to have*. Williams has often been read as intending a sense of possibility defined in terms of roughly this sense of rationality; *i. e.*, S1 claims that R is a reason for A to ϕ only if A

4. Some passages suggest a reading toward the other end of the spectrum; "If something can be a reason for action, then it could be someone's reason for acting on a particular occasion" (p. 106). This could be read as offering the weak condition that there be some way of being an agent, realized by some actual individual, such that if A was like that, he would be motivated to ϕ by acceptance of R . Interpreted this way (S1) is quite plausible. But S2 now preposterously claims that the motivational possibilities for any agent are constrained by the actual motivational set of A .

would be motivated to ϕ by believing R if he was ideally rational. Many philosophers consider this claim a “platitude”,⁵ even many who Williams would apparently consider defenders of “external reasons”. I call this a *rational possibility* reading of the argument.

There is evidence that by ‘rationality’ Williams means something narrower, however, and hence intends a different sense of possibility. He is concerned with what would motivate the actual agent if he were to deliberate “rationally”, and equates “rational deliberation” with *fully imaginative and valid thought proceeding from a belief-set purged of errors*. For economy I call this *sound* deliberation, as Williams does. His focus on sound deliberation can be and has been seen as simply resulting from his substantive view of what rationality in the broad sense amounts to, but there is some basis for taking his concern with sound deliberation to be independent of the platitude about rationality. S1 would then connect an agent’s reasons with what he can be motivated to do by sound deliberation from his actual motivational set. I call this a *deliberative possibility* reading.

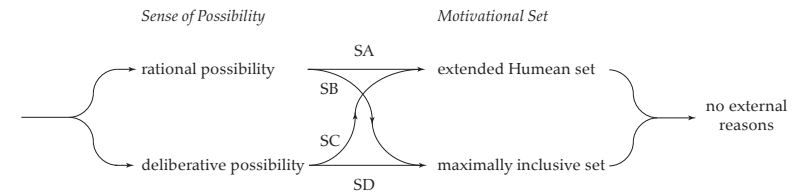
(ii) The second interpretive question concerns Williams’ notion of a “motivational set”, which as I observed also concerns what he means by “external reasons” and hence the true target of his skepticism. What psychological elements are possible members of a “motivational set”? Although he sometimes uses ‘desire’ as a term of art for any member, he explicitly disavows the “sub-Humean model” which claims that all actions are motivated by desires proper, and he professes a more “liberal” view. Just how liberal his view might be is controversial. On one common interpretation it remains essentially Humean; the many different elements that can comprise a motivational set may not all be “desires” strictly speaking, but are all desire-like, being pro-attitudes seeking *satisfaction* of their representational contents. Many commentators therefore ascribe a “desire-based” theory of reasons to Williams despite his disclaimers.⁶ Call this an *extended Humean set* reading of the argument.

5. e.g. Nagel 1970; Darwall 1983; Korsgaard 1986; Wallace 1999.

6. Wallace 1990: 375n, 376 and 1999: 218; Cohon 1986, 1993; Millgram 1996; Shafer-Landau 2003: 172–5.

A rival reading takes Williams at his word, allowing motivational sets to extend as widely as possible, including anything in an agent’s psychological state that is relevant to what he can be motivated to do. The essential issue is the place of *dispositions*; are they in or out? On the broadest interpretation, the motivational set includes mere dispositions to be motivated to act, together with the desire-like elements.⁷ Call this a *maximally inclusive set* reading of the argument.

These two rival interpretations of the motivational set combine with the two rival interpretations of the sense of possibility to yield four different versions of the Standard Argument, as follows:



Cases have been made for attributing each of these four versions of the argument, SA, SB, SC, and SD, to IER. I shall now examine each in turn, observing its flaws. I then show that there are serious exegetical obstacles to reading any of these into Williams’ paper.

Version SA: the rational instrumentalist argument⁸

SA1. R is a reason for A to ϕ only if A would, if ideally rational, be motivated to ϕ by believing R ;

7. The distinction between *dispositional desires* and dispositions to *have* desires complicates the issue. An enhanced Humean set can include the former (accommodating Williams’ inclusion of “dispositions of evaluation”), but not the latter (see Millgram 1996: 208–9). Dispositional analyses of desire complicate matters further (Smith 1994; Joyce 2001: 111; see discussion in Wallace 1990: 359–60). Proponents of these analyses will deny a distinction between enhanced Humean and maximally inclusive versions, and deflate the former into the latter.
8. E.g., Cohon 1986: 553 and 1993: 272; Hooker 1987; Wallace 1990: 376; Smith 1995; Velleman 1996; Millgram 1996: 210–11; Joyce 2001: 132; Shafer-Landau 2003: 172; Setiya 2004: 268–9; Pettit & Smith 2006: 145; Brunero 2007: 24.

SA2. *A* would, if ideally rational, be motivated to ϕ by believing *R* only if *A* has a desire-like state that could bring *A* to be motivated to ϕ by believing *R*;

SA3. If *R* is an external reason for *A* to ϕ , then *A* does not have a desire-like state that could bring *A* to be motivated to ϕ by believing *R*;

SA4. Therefore there are no external reasons.

This most common interpretation combines the rational possibility and enhanced Humean set readings. The argument begins from the platitude SA1 that reasons must motivate an agent insofar as he is rational. SA2 then introduces a thesis about what it is to be rational: that the motivations of the *rational A* are constrained by the desire-like states of the *actual A*.⁹

But should we accept this thesis about rationality? Perhaps if *A* were rational, he would have some desire that he actually lacks, which would bring him to be motivated to ϕ . Perhaps *A* isn't rational precisely *because* he lacks such an element in his motivational set, like a concern for others, an aversion to pain, or a desire to act on his reasons.¹⁰ Indeed Williams himself grants that if *A* were rational, his motivational set may include elements that it actually lacks. But he restricts these elements to motivations that can be derived from *A*'s existing motivations.

Williams would then be assuming a particular conception of the nature and powers of practical rationality: the "Humean" or *instrumentalist* view that practical reason cannot by itself produce or entail a motive, but merely facilitates flow of motivational force from pre-existing desires, channeling it from one object to another. "At least one contemporary philosopher, Bernard Williams"; Christine Korsgaard

9. The claim is not that the motivations of *any* rational version of *A* must be restricted by *A*'s actual desires, only that in the closest world in which *A* is rational *A*'s motivations are restricted by *A*'s actual desires. Rationality *by itself* doesn't require one to have or lack any particular motivational element.

10. Korsgaard 1986; Smith 1995: 118; Velleman 1996: 696–7; Parfit 1997: 101.

writes, "has taken something like Hume's argument [that all rational motivation must ultimately spring from some nonrational source] to have this kind of independent force, and has so argued in his essay 'Internal and External Reasons'" (1986: 8). But, Korsgaard objects, this "motivational skepticism" about practical reason actually presupposes "content skepticism" — that reason by itself has *nothing to say* about what motives to have. It assumes a narrow, *procedural* view of rationality, as merely discovering various relations and matters of fact; as we've observed, Williams equates "rationality" with sound deliberation. The procedural view of rationality yields in turn the instrumentalist view of practical rationality, as reasoning to the means to satisfying pre-existing desires.

It is question-begging, however, to argue against external reasons from an assumption of a procedural or instrumentalist conception of rationality. Given the platitude that an agent is rational only if he is disposed to be motivated by his reasons, someone who accepts the existence of external reasons will believe that an agent is not rational unless he is disposed to be motivated by these external reasons, and will favor a *substantive* rather than a procedural account of rationality: that being rational entails having certain elements (dispositions, desires, etc.) in one's motivational set. John McDowell thus objects to Williams that becoming rational (or, as he prefers, virtuous)¹¹ might be a process of *conversion* rather than of sound reasoning (1995), and Elijah Millgram suggests that it might require having certain experiences (1996: 202). Williams is thus widely accused of *begging the question regarding the nature of rationality*.¹²

11. Some commentators reject the platitude linking responsiveness to reasons with "rationality", preferring to link it with virtue, but here we can ignore this distinction. These critics are often willing to grant Williams a procedural view of "rationality" (McDowell 1995; Millgram 1996; Shafer-Landau 2003; Setiya 2004, 2007), but accordingly they give the platitude a wider reading: an agent is only virtuous (a "correct deliberator") if he is motivated by his reasons, and what it is to be virtuous doesn't depend on one's actual motivational set. Sobel (2001a) rather defends Williams' argument on the wider reading.

12. Korsgaard 1986; Hooker 1987; Wallace 1990; Cohon 1993: 272; Parfit 1997: 116; Pettit & Smith 2006: 151–2.

Version SB: *the modest internalist argument*¹³

SB1. *R* is a reason for *A* to ϕ only if *A* would, if ideally rational, be motivated to ϕ by believing *R*;

SB2. *A* would, if ideally rational, be motivated to ϕ by believing *R* only if *A* has at least some disposition to be motivated to ϕ by believing *R*;

SB3. If *R* is an external reason for *A* to ϕ , then *A* does not have even a disposition to be motivated to ϕ by believing *R*;

SB4. Therefore there are no external reasons.

This second version combines the rational possibility and maximally inclusive set readings. SB1 is the same platitude as before, connecting reasons and rational motivation, but SB2 eschews the question-begging instrumentalist theory of rationality for a much weaker claim linking the rational possibility of motivation by *R* with *A*'s having some disposition to be motivated by *R*. This is the weakest of the four versions of the Standard Argument, and it is compatible with neo-Kantian views like Korsgaard's which claim that pure Reason can motivate by itself and without desire.¹⁴ Hence Korsgaard writes, in endorsement of SB2,

If one accepts the internalism requirement, it follows that pure practical reason will exist if and only if we are ca-

13. Cohon 1986: 555n; Parfit 1997: 118–9; Scanlon 1998: 369.

14. It is still stronger than some versions of "internalism" about reasons, which claim merely that *R* can be a reason for *A* to ϕ only if *A* would, if he were ideally rational, have a disposition to be motivated by *R* to ϕ — *i. e.*, tying reasons only to the *rational* agent's maximally inclusive motivational set (*e.g.* Smith 1995: 117). Hooker (1987) suggests this is a possible interpretation of Williams' internalist premise. Cohon (1993: 285–6n4) suggests that Williams is mistaken to claim agreement with Korsgaard, because she accepts only the weaker internalism. But Korsgaard doesn't believe that anything could be a reason for *A* to act if *A* lacked the capacity to reason, and this capacity qualifies for membership in the maximally inclusive set (Cf. Velleman 1996: 697n). This misreading of Korsgaard is presumably due to the ambiguity of 'rational agent', by which she sometimes means an agent merely with a rational faculty.

pable of being motivated by the conclusions of the operations of pure practical reason as such. Something in us must make us capable of being motivated by them, and this something will be part of the subjective motivational set. (1986: 21; also 24)

Williams' only real disagreement with Korsgaard would then be over an issue external to his internal reasons thesis: whether rationality involves substantive rather than merely procedural requirements. But this version allows that if Korsgaard is right about that, then Kantian reasons are internal. Indeed, Williams later writes that "in this respect, I have no basic disagreement with [her] excellent paper" (1995a: 44n; also 2001: 94, 97n4).

While this version's characterization of rationality in SB2 may have wider appeal than instrumentalism, it is no less question-begging against external reasons¹⁵ — as the external reasons thesis is identified by SB3. The external reasons theorist is here the (somewhat rarer) person who thinks that agents need not have any disposition to be motivated by their reasons. Given that SB1 is a platitude, if we are unsure whether all reasons are internal we should also be unsure whether the ideally rational agent's motivations are constrained by the dispositions of the actual agent. Therefore it seems that we would only find SB2 compelling if we had already come to the conclusion that there are no external reasons (SB4).

Version SC: *the Humean motivation argument*¹⁶

SC1. *R* is a reason for *A* to ϕ only if *A* would, if he deliberated soundly, be motivated to ϕ by believing *R*;

SC2. *A* would, if he deliberated soundly, be motivated to ϕ by believing *R* only if *A* has a desire-like state that could bring *A* to be motivated to ϕ by believing *R*;

15. Shafer-Landau 2003: 179.

16. See, *e.g.*, Cohon 1986 and 1993: 266; Hollis 1989: 173; Parfit 1997: 112; Sobel 2001a; Joyce 2001: 110; Heuer 2004: 46; Setiya 2007: 102.

SC3. If R is an external reason for A to ϕ , then A does not have a desire-like state that could bring A to be motivated to ϕ by believing R ;

SC4. Therefore there are no external reasons.

This version combines the deliberative possibility and enhanced Humean set readings. Williams' argument then wouldn't begin with the platitude connecting reasons with rationality, but rather with the weaker claim (SC1) that R is a reason for an agent to ϕ only if being motivated to ϕ by believing R is *deliberatively* possible for him. SC2 applies a neo-Humean theory of motivation: believing R would motivate A *post*-deliberation only if there was some desire-like element present in A 's *pre*-deliberation motivational set that would bring A to be motivated to ϕ by believing R . By SC3, external reasons don't meet this condition: "For, *ex hypothesi*, there is no motivation for the agent to deliberate *from*, to reach this new motivation" (p. 109). *Ergo* there are no external reasons, *i. e.*, no reasons that have no motivational connections with the agent's existing desires.

A first problem here is that when we replace the rational possibility in the previous versions with mere deliberative possibility, we trade a reading of S1 which is widely considered to have the indubitability of a platitude for a reading (SC1) which seems unmotivated and un compelling. It is a widespread and commonsense view, for example, that the existence of an agent's moral and prudential reasons don't depend on what he can be motivated to do merely by sound deliberation. So why accept that R could be a reason for A only if he would be motivated by R after sound deliberation? Several commentators argue that we can even have desire-based reasons that sound deliberation could not possibly bring us to be motivated by.¹⁷ Williams' rationale for SC1

17. Cohon 1993; Millgram 1996; Johnson 1999; Sobel 2001b; Shafer-Landau 2003: 186. These are desire-based reasons that an agent has because of his incapacity to conduct sound deliberation. If such reasons exist it could not be the case that sound deliberation could bring us to be motivated by them, because if we were able to carry out that sound deliberation then the reasons would not have existed.

would apparently be the "explanatory dimension" of reasons statements — that we can explain agents' actions by citing their reasons for acting (and, therefore, that their reasons must have some connection to what motivates them). But this is an inadequate rationale for the premise. We might be able to explain an agent's actions by appeal to his reasons even if agents occasionally can't be motivated by some of their reasons: perhaps we assume the presence of a contingent desire to act on one's reasons that is sometimes silenced, for example. Or we might deny that it is essential to *normative* reasons, as opposed to "motivating" or "explanatory" reasons, that they be explanatory of *deliberatively* possible actions.¹⁸

Version SC faces a more serious problem in SC2. This premise articulates the neo-Humean theory of motivation, that all motivation derives from desire-like psychological states. This theory's proponents often treat it as an *a priori* truth. But motivation is a causal process and (as Hume himself insists) causal relations can presumably be known only empirically.¹⁹ As Jay Wallace writes,

One is initially tempted to respond on behalf of the externalist with a reminder that there are no *a priori* constraints on causal relations. If the question is, how could reflection by itself produce a new motivation to action?, the answer might simply be that there is no general reason to suppose that it could not. (1999: 220–21)

The neo-Humean theory has not been empirically confirmed. Perhaps the *belief that R*, or the *belief that R is a reason to ϕ* , might themselves be sufficient sources of motivation for A to ϕ .²⁰ In that case, R could meet the definition of external reasons in SC3 while SC2 is false. That is to

18. For the distinction, see Nagel 1970; Smith 1994. It is however vexed: see Darwall 1983; Parfit 1997: 114n; Dancy 2000; Finlay 2006, 2007.

19. Causal connections can be *a priori* when the cause and effect are interdefined. I have myself argued for an *a priori* link between desire and motivation (Finlay 2007; see also Mele 2003: 53; Smith 1994; Wallace 1990: 359–61). But an argument is necessary, and Williams doesn't provide one.

20. Nagel 1970; Darwall 1983; Wallace 1990: 365–6, 376–8 and 1999: 228; Cohon

say, the belief that R , which is *not* in the agent's motivational set prior to deliberation, might be a sufficient cause of motivation to ϕ . Therefore, contra SC2, A can be motivated through deliberation by coming to believe R even though, prior to deliberation, his motivational set contains nothing that plays a role in producing that motivation. Deliberation produces the motivation simply by producing the belief. Williams is therefore commonly accused of dogmatically assuming a faulty theory of motivation.²¹

Even if the neo-Humean theory of motivation is true, SC would still seem to be fallacious. Granting that desire is a necessary source of any motivation, it doesn't follow that A could be motivated through deliberation only by virtue of some desire he had *prior* to deliberating. The process of deliberating itself could stimulate new, nonderivative desires, which could then motivate A to ϕ .²²

*Version SD: the incoherence argument*²³

SD1. R is a reason for A to ϕ only if A would, if he deliberated soundly, be motivated to ϕ by believing R ;

SD2. A would, if he deliberated soundly, be motivated to ϕ by believing R only if A has at least a disposition to be motivated to ϕ by believing R ;

SD3. If R is an external reason for A to ϕ , then A does not have even a disposition to be motivated to ϕ by believing R ;

SD4. Therefore there are no external reasons.

1993: 270; Velleman 1996: 699; Parfit 1997: 105; Dancy 2000; Shafer-Landau 2003.

21. Cohon 1986: 549–50; Wallace 1990, 1999; Shafer-Landau 2003; Parfit 1997: 112, 115; Setiya 2007: 102.

22. This point is made in Nagel 1970, Darwall 1983, and pressed against Williams in Wallace 1990: 376; Cohon 1993: 271; Smith 1995: 115; Parfit 1997: 105–6; Shafer-Landau 2003: 175.

23. Millgram 1996: 198.

This fourth and final version of the Standard Argument combines the deliberative possibility and maximally inclusive set readings. It therefore shares the first premise of SC, which we observed to be unmotivated and unconvincing. But rather than the problematic neo-Humean theory of motivation, SD2 claims the extremely weak thesis that an agent can be motivated to ϕ by believing R through sound deliberation only if he has *some* disposition in virtue of which he might be motivated to ϕ by believing R . It therefore avoids the charge of assuming a faulty theory of motivation; indeed SD2 seems trivially true, and Williams writes, “certainly — and nobody denies this — what he actually does has to be explained by his [motivational set]” (1989: 39). If the belief that R (or that R is a reason to ϕ) is sufficient to motivate A to ϕ through sound deliberation, then R is in fact an *internal* reason for A to ϕ ;

[T]his agent, with this belief, appears to be one about whom, now, an *internal* reason statement could truly be made: he is one with an appropriate motivation in his [motivational set]. A man who does believe that considerations of family honour constitute reasons for action is a man with a certain disposition to action.... (p. 107)

Therefore, in Millgram's words, “would-be external reason explanations turn out, when they are successful, to be *internal* reason explanations” (1996: 198).²⁴

Like SB, this argument appears to beg the question against external reasons. It provides the same definition of an “external reason” (SB3), as something by which the agent has no disposition to be motivated. But since SD2 is trivial (and Williams apparently thinks that “nobody denies it”), SD1 is effectively just the claim that R is a reason for A only if he has a disposition to be motivated by it. In other words, the first premise SD1 — which seems unmotivated and unconvincing — essentially claims that there are no “external reasons”, thereby begging the

24. Also Millgram 1996: 207–8; Shafer-Landau 2003: 179; Brunero 2007: 29.

question in the most blatant way. Hence Russ Shafer-Landau objects that “whether having a reason entails the possibility of being motivated upon [sound] deliberation is precisely what is at issue” (2003: 179).²⁵

So far I have sketched four versions of the Standard argument that can and have been attributed to Williams, and I have detailed how weak (question-begging or unconvincing) each is known to be. These observations do not yet amount to a decisive case against reading Williams in any of these ways. So let us turn to the exegetical issue: which of these versions can plausibly be read into the text? We shall see that there are significant obstacles to any of the interpretations, and indeed to rational *or* deliberative possibility readings, enhanced Humean *or* maximally inclusive set readings. These obstacles fall into three categories. (1) Violations of charity: we shouldn’t adopt readings that attribute particularly crazy, naïve, or obviously false views to Williams if we can avoid it; (2) Inconsistency: we should avoid readings on which his claims are incompatible or contradictory; (3) Author’s authority: as I’ve observed, Williams later offers defenses and clarifications of his arguments in IER, and elsewhere explores some of the same issues. These should not be discounted lightly. While some commentators suggest that Williams may or must have changed his mind about parts of his original argument, we should note that he never expresses any reservations about the argument or offers any (except terminological) retractions. Instead, he seems to defend his original argument against the criticism.

The rational instrumentalist (SA) and modest internalist (SB) versions have in common the platitudinous rational possibility reading of premises S1 and S2. Charity aside, the problem with reading Williams this way is that he later explicitly repudiates this interpretation, agreeing with McDowell that “it was a mistake to pick out ‘irrational’ as a

25. Shafer-Landau writes “rational deliberation” here, and calls this “rational possibility”. However, he seems to mean rather what I am calling “deliberative possibility”, since he restricts it to motivation through sound deliberation, and contrasts it with the possibility of being motivated if *substantively* rational (2003: 173–4).

crucial term in this connection” (1995: 192; also see 2001: 93). Some commentators interpret this rather as a *retraction*,²⁶ but this is in tension with what Williams says next; “But I want still to press the point that I made originally, that ... it is important that [the reasons statement] ... should say something special about *A*, and not merely invoke in connection with him some general normative judgment.” This suggests Williams never intended to base his argument on a claim about how agents *ought* to be. Further evidence is his emphasis, in IER, on the necessary “explanatory dimension” of reasons statements; he claims that *R* can be a reason for *A* to ϕ only if *R* could figure in an explanation of why the *actual A* would ϕ , and not rather why some idealized counterpart of *A* would ϕ .

Rational possibility readings are also in tension with Williams’ repeated claims that “there is an essential indeterminacy in what can be counted a rational deliberative process” (p. 110; also 1989: 38; 2001: 92). He suggests that this places an “onus of proof” on the external reasons theorist rather than on him, but as Rachel Cohon points out (1986: 555), we need a substantive idea of what counts as rational deliberation before we can assess any version of the premise S2, which claims that the possible outputs of rational deliberation are limited by the agent’s motivational set.²⁷

The other two interpretations, the Humean motivation (SC) and incoherence (SD) versions, involve deliberative possibility readings of S1 and S2. These also have problems, despite the strength of the case against the rival, rational possibility readings. First, Williams recognizes ways in which coming to believe an external reasons statement might motivate an agent, but he discounts them as irrelevant, because they are improper ways of becoming motivated (p. 108–9). This is surely a normative rather than a purely explanatory constraint, which may seem in tension with deliberative possibility readings.²⁸

26. Setiya 2004: 269.

27. Also Korsgaard 1986: 23; Hollis 1989: 174.

28. His notion of sound deliberation involves some idealization, however, so perhaps this objection is not so worrisome. But if the basic rationale for Williams’

A second problem comes from the apparently unmotivated and un-compelling character of S₁ on these readings. Hence Millgram holds that an instrumentalist account of rationality — an SA reading — “is the extra ingredient that makes internalism about reasons a reasonably held position”, and claims that this is a “controlling exegetical consideration, one which underwrites an instrumentalist reading at ambiguous moments in Williams’ writing” (1996: 213).²⁹ Charity weighs against attributing to Williams an argument based on a premise as un-compelling as SC₁ and SD₁ appear to be. Some commentators offer the uncharitable explanation that Williams simply overlooked the distinction between normative and explanatory reasons, which Nagel had already drawn in 1970.³⁰ But this seems unlikely, as Williams warns against the idea that “the internal reason conception is only concerned with explanation, and not at all with the agent’s rationality” (p. 103), and discusses the distinction in his later revisitations of the argument (1989: 39; 2001: 93). It is therefore puzzling why Williams would have built his argument on a version of S₁ involving merely the deliberative possibility of motivation.

The choice between the rival interpretations of the motivational set presents more serious difficulties. Most commentators favor an enhanced Humean set reading of the argument, SA or SC.³¹ There are numerous problems here:

(i) Williams *ought* to have known better than to assume — without argument — that actual or rational motivation requires antecedent desire-like states. Beside the dogmatism of the view, Nagel’s influential case against it (1970) had appeared nine years earlier — a point on which Williams is admonished by Pettit & Smith (2006: 152). More

internalism is the “explanatory dimension”, then it is something of a puzzle why there is idealization in his account at all.

29. Also Shafer-Landau 2003: 179; Skorupski 2007: 76–9.

30. Parfit 1997: 113; Shafer-Landau 2003: 178.

31. Millgram correctly points out that Williams’ examples of deliberation often involve the *satisfaction* of elements in the motivational set (1996: 212), but this is compatible with their merely being paradigms.

significantly, there is evidence that he *did* know better. (We shall see that plausibly IER is partly a response to Nagel).

(ii) This reading is forced to disregard or downplay Williams’ explicit rejection of a “sub-Humean”, instrumentalist model in favor of a “liberal” view (p. 104; also 2001: 92). Hence Millgram admits the “author’s likely surprise at having the label [of ‘instrumentalist’] adhere to his own carefully nuanced views” (1996: 213).

(iii) He explicitly considers and grants suggestions about how agents can be motivated that look *prima facie* incompatible with this reading: (a) that a belief may by itself motivate *A* to ϕ (p. 107); (b) that exercise of imagination can generate new, underived desires and motivations (p. 105); (c) that if an agent has a disposition to be motivated to ϕ by some belief, then this disposition is in his motivational set, and he has an internal reason to ϕ (p. 107; also 1989: 37).

(iv) In later writings, he clearly rejects a Humean reading of S₂. He expresses “basic agreement” with Korsgaard’s claim that if all agents have a faculty of practical reason which when properly exercised would lead them to be motivated by some *R*, then *R* is an internal reason for everybody, and he attributes to Kant the view that there are only internal reasons (1985: 223–4).³² He further claims that “nobody denies” that what an agent “actually does has to be explained by his [motivational set]” (1989: 39) — although he can’t have been unaware of the growing philosophical opposition to the neo-Humean theory of motivation.

This all adds up to strong evidence against the enhanced Humean set readings SA and SC, and might be thought a strong case by elimination for the weaker, maximally inclusive set readings, *i.e.*, the modest internalist (SB) or incoherence (SD) versions. Unfortunately, a maximally inclusive set reading also faces some serious exegetical difficulties, leading Jay Wallace to write that “it is difficult to reconcile with the text ... which connects ... with the Humean approach” (1990: 375n). A first problem is that as we’ve already noted, SB and SD are

32. Wallace (1990: 375n) suggests that this represents a change in view.

blatantly question-begging, which is an obstacle to reading Williams charitably in this way.

Second, Williams' characterization of deliberation as proceeding *from* elements in the motivational set presents a problem. This seems to presuppose some broadly Humean account of the motivational set, and to exclude mere dispositions. We do not deliberate *from* our dispositions to be motivated, to corresponding motivation.³³

A third difficulty is Williams' apparent view of his argument's dialectical significance. These versions share an interpretation of S₃ that effectively defines external reasons as reasons that the agent has no disposition to be motivated by. Some commentators have thus understood Williams' basic point against "external reasons theorists" to be that agents cannot have reasons that are powerless to motivate them.³⁴ But this is apparently not what *Williams* thinks is at issue between his opponents and himself. Unlike other "internalists" such as Korsgaard and Stephen Darwall, Williams assumes (at least in IER) that external reasons theorists agree that it is — and maybe even that it has to be — possible for an agent to be motivated by an external reason. He writes, "what has to hold for external reasons statements to be true ... is that the new motivation could be in some way rationally arrived at, granted the earlier motivations" (p. 109), and "the external reasons theorist essentially wants, that the agent should acquire the motivation *because* he comes to believe the reason statement" (p. 108–9; see also 1989: 39; 1995: 189). That external reasons beliefs can motivate is, in fact, a point that he "grants" to his opponents (p. 107).³⁵ This is a serious obstacle to a maximally inclusive set reading, SB or SD, of

Williams' argument. Hence Derek Parfit rejects maximally inclusive set readings, writing,³⁶

Williams would not, I believe, give this reply [that if we came to be motivated, we must have had a disposition to be motivated]. It would achieve nothing. No Externalist would mind conceding that, if our deliberation leads us to make some decision, we must have been such that our deliberation might lead us to make this decision. (1997: 119)

To summarize: (1) the rational possibility readings SA and SB conflict with Williams' explicit statements about the sense and role of "rationality" in his argument. But (2) the deliberative possibility readings SC and SD seem weaker and less plausible arguments, since their interpretation of S₁ seems unmotivated and unconvincing. (3) The enhanced Humean set readings SA and SC conflict with Williams' many explicit claims supporting a maximally inclusive account of a motivational set, and would have him embracing either a question-begging account of rationality (SA), or a faulty theory of motivation (SC). (4) But the maximally inclusive set readings SB and SD are incompatible with Williams' account of deliberation as proceeding from existing motivations, and with Williams' own understanding of what is at issue between himself and his opponents. Every reading is therefore at least doubly problematic.

One way of trying to deal with some of the inconsistencies and interpretive problems is to suggest that Williams offers more than one of these arguments.³⁷ Millgram interprets him as giving roughly the incoherence argument SD ("from explanation") on pp. 102–7, and the rational instrumentalist argument SA ("from deliberation") on p. 109 (1996: 197–8). This proposal faces several problems. Williams himself doesn't distinguish between an argument from explanation and an argument from deliberation, and his discussions of explanation and de-

33. Millgram 1996: 208–9; Parfit 1997: 115; Scanlon 1998: 367–9.

34. Parfit 1997: 100; Scanlon 1998: 369; Joyce 2001: 108; Shafer-Landau 2003: 179.

35. Might Williams' thought be that the external reasons theorist can postulate the existence of a disposition to be motivated by the belief that *R* is a reason, in the absence of any disposition to be motivated by the belief that *R*? This wouldn't rescue the reading, however, as he also writes that "the external reasons statement will have to be taken as ... entailing the claim that if the agent rationally deliberated, then ... he would come to be motivated to ϕ " (p. 109).

36. Also Cohon 1986: 555n; Wallace 1999: 221; Setiya 2004: 269.

37. Also Parfit 1997: 112; Skorupski 2007: 87.

liberation are intertwined throughout IER. He introduces his account of deliberation as early as p. 102, and he is still talking about explanation on p. 109. Further, these two arguments involve incompatible interpretations of the notion of a “motivational set”; SA involves an enhanced Humean set reading, and SD a maximally inclusive set reading. Williams would then not only be offering two weak arguments, but offering arguments that require incompatible interpretations of their shared terms. Other commentators accordingly suggest that Williams equivocates, sliding between different senses of ‘motivational set’ and ‘possibility of motivation’.³⁸ Still others claim that his argument changes over time.³⁹

One final exegetical puzzle needs our attention: Williams’ challenge to external reasons theorists to identify the *conceptual content* of an external reasons claim, on pp. 108–110. As he later puts it,

What is it that the agent comes to believe when he comes to believe he has a reason to ϕ ? If he becomes persuaded of this supposed external truth ... what is it that he has come to believe? This question poses a challenge to the externalist theorist. (1989: 39)

The four versions of the Standard Argument considered above turn on claims merely about the necessary conditions for having a reason, not on what reasons claims mean, and so this challenge is generally identified as a distinct argument. Millgram describes it as “one of Williams’ subsidiary arguments” (1996: 201). But it is hard to fit this “conceptual content argument” into the dialectic of IER in any charitable way.

Williams presses the conceptual challenge to external reasons on p. 108, after considering what is apparently the *faulty theory of motivation* objection, that an agent might be motivated merely by his belief that he has an external reason. He then writes, “this agent, with this belief, appears to be one about whom, now, an *internal* reason statement

could truly be made: he is one with an appropriate motivation in his [motivational set]” (p. 107). On SB and SD this would be the conclusion of the Standard Argument; that (as Millgram puts it) “would-be external reason explanations turn out, when they are successful, to be *internal* reason explanations” (1996: 198). But he doesn’t continue as if he had just proven his case, instead raising the question about conceptual content: “Now it does not follow from this that there is nothing in external reasons statements. What does follow is that their content is not going to be revealed by considering merely the state of one who believes such a statement ...” (pp. 107–8).

This move is apparently repeated a page later. He articulates again an objection to the Standard argument (either the *faulty theory of motivation* or the *question-begging account of rationality* objection):

It might be said that the force of an external reason statement can be explained in the following way... . [A] rational agent is precisely one who has a general disposition in his [motivational set] to do what (he believes) there is reason for him to do. So when he comes to believe that there is reason for him to ϕ , he is motivated to ϕ , even though, before, he neither had a motive to ϕ , nor any motive related to ϕ -ing in one of the ways considered in the account of deliberation. (p. 109)

This time he doesn’t suggest that a disposition of this kind would entail that the agent had an internal reason. Instead, he immediately issues the conceptual content challenge, to which he suggests an answer: “*What* is it that one comes to believe when he comes to believe that there is reason for him to ϕ , if it is not the proposition ... that if he deliberated rationally, he would be motivated to act appropriately?” But, he goes on to conclude, in that case there presumably are no external reasons.

This seems a woefully ineffectual response to the objection. First, Williams proposes a content for external reasons claims that is seemingly plucked out of thin air, without argument. It is mysterious why

38. Cohon 1986: 552; Scanlon 1998: 367–9; Shafer-Landau 2003: 178–9.

39. Wallace 1990: 375n; Setiya 2004: 269.

they would have to mean *that*, to be potentially motivating.⁴⁰ For any plausible analysis of external reasons beliefs there is a possible disposition to be motivated by beliefs of that kind (and an arguable theory that rationality entails such a disposition). Hence Wallace responds,

Externalists would agree that the basal desire they appeal to cannot fix the content for our beliefs about what we have reason to do. But they would insist that this is not the role assigned to that desire. Its job is not to fix the content of our reasons, but to explain how beliefs about our reasons can generate a corresponding motive. (1999: 228n)

To Williams' rhetorical question about what *else* such statements might mean, his opponent might serenely suggest whatever content she likes, or like many today, simply deny the possibility of informative analysis of reasons-beliefs.⁴¹ The external reasons theorist might even accept Williams' suggested analysis, because it isn't obvious that an agent couldn't be disposed to be motivated to ϕ by believing that he would be motivated to ϕ if he were to deliberate rationally. Michael Smith (1994) suggests that a very similar disposition is essential to being rational, for example.

Why then would Williams have thought that this conceptual content challenge had any force against the objections to the Standard Argument? There is some indication that his point is not that these beliefs could not motivate, but rather that they could not be *true*: "We were asking how any true proposition could have that content; it cannot help, in answering that, to appeal to a supposed desire which is activated by a belief which has that very content" (pp. 109–110). But (i)

40. Folke Tersman observed to me that if Williams is concerned with what external reasons beliefs must mean in order to be *necessarily motivating for rational agents*, then he might reasonably suppose that only this content could fit the bill. However, Williams would still be leaping to an insufficiently argued conclusion, and assuming a question-begging account of the nature of rationality.

41. Cohon 1986: 553; Wallace 1990: 377; Millgram 1996: 207; Parfit 1997: 109–10, 120–1.

this seems to commit a straw man fallacy, since the external reasons theorist is apparently free to reject Williams' proposed analysis, and (ii) it seems viciously circular. *Why* couldn't it be true that rational or sound deliberation would motivate the agent "whatever motivations he originally had" (p. 109)? Presumably he is here appealing to the restrictions the Standard Argument places on possible motivation. But as we've seen, Williams seems to raise this "subsidiary" conceptual content argument as a response to objections to the Standard Argument calling into question precisely these restrictions. They cannot therefore be used in turn to support the conceptual content argument.

Matters look no better if we rather suppose, as some have, that the conceptual content argument is intended as an entirely separate argument against external reasons. The argument would still suffer from the other flaws we've noted, and Williams would then be committing an elementary dialectical sin, reacting to a serious objection to a premise of his main argument, not by defending it, but by offering an entirely different argument to support the conclusion.

In sum, if Williams intended to advance one or more of the versions of the Standard Argument, then his arguments seem weak and his position seems confused and inconsistent. Further, since he continued to stand behind his argument(s) over the next twenty-plus years in the face of all these objections and problems, he would have to have been intransigently obtuse. Many commentators confess uncertainty about what Williams really meant, and the difficulty of interpreting his arguments is often remarked.⁴² This might be attributed to the uncommonly muddled and unclear nature of his writing in IER, but we might rather wonder whether these commentators hold mistaken preconceptions about the intended argument, and are struggling to cram the text into a mould foreign to its design. Williams himself later observed that "a number of discussions has [sic] led me to think there is something about this view, or the ways I have so far found to express it, that easily leads to misunderstanding" (1989: 35). To ascribe Williams

42. Wallace 1990: 375, 377; Millgram 1996: 199, 214; Velleman 1996: 700n; Parfit 1997: 116n; Pettit & Smith 2006: 142; Brunero 2007: 29.

a version of the Standard Argument would therefore be grossly uncharitable if there is any exegetically defensible or preferable reading on which his argument is stronger.

2. The Alternative (or Actual) Argument

There is an alternative way of reading Williams' argument, one which (I will argue) is compatible with virtually everything in IER and subsequent texts, and which isn't vulnerable to *any* of the objections and exegetical difficulties catalogued above. On this interpretation, Williams offers just one consistent argument, which unites his central concerns with (i) the explanatory dimension of reasons statements, (ii) their conceptual content, and (iii) the connection between reasons and deliberation.

Williams' concern with explanation is uncontroversially pivotal for his argument;⁴³ on p. 102 he writes,

[T]his explanatory dimension is very important, and we shall come back to it more than once. If there are reasons for action, it must be that people sometimes act for those reasons, and if they do, their reasons must figure in some correct explanation of their action.

Commentators have assumed he is here merely asserting the common thesis that for a consideration to be a reason for an agent, it must have the capacity to motivate that agent (whether this is rational or deliberative possibility),⁴⁴ and hence that we can cite facts about reasons as part of an explanation of agents' actions. Korsgaard writes that Williams' point is "that unless reasons are motives, they cannot prompt or explain actions" (1986: 11). My hypothesis is that Williams holds a much more radical and interesting view: that the very conceptual content of "a reason for action" involves the concept of an explanation. To

43. In 1989 he tells us that "the first [fundamental motivation of the internalist account] is the interrelation of explanatory and normative reasons" (p. 38).

44. Nagel 1970; Korsgaard 1986; Wallace 1999; Darwall 1983.

believe that *R* is a reason for action is *nothing other* than to believe it is a special kind of explanation of action.

Williams' focus on the conceptual content of reasons statements does not begin with the appearance of the supposed "conceptual content argument" on pp. 108–9; rather it is ubiquitous throughout IER and also his later revisitations of his argument ("What is it the agent comes to believe when he comes to believe that he has a reason to ϕ ?" [1989: 39]). He *begins* in IER by contrasting two different "interpretations" or "senses" of reasons statements (p. 101). Interpreting the "explanatory dimension" as a feature of the meaning of reasons statements is therefore reasonable. By contrast, no version of the Standard Argument requires any claims about the meaning of reasons statements; *S*₁ merely offers a necessary condition for *R*'s being a reason. Williams' conceptual focus would therefore be gratuitous on the standard readings.

What Williams explicitly says about the "explanatory dimension" does not decisively support this conceptual interpretation over a more orthodox reading. The best evidence for the conceptual interpretation is that it makes the most sense of Williams' entire argument and other claims, as I shall show. We can motivate this interpretation independently by considering the historical context, however. Most discussions of reasons prior to the publication of IER in 1979 did not draw sharply the now common distinction between "normative" and "motivating" (or "explanatory") reasons, and a dominant view was that a single concept of a reason for action somehow filled both roles. Donald Davidson's influential paper "Actions, Reasons and Causes" (1963) argued that the concept of "reasons" that justify an action just is a certain concept of "reasons" that causally explain it; a justification is simply a *rationalizing* explanation in terms of the agent's desires and beliefs. What could we plausibly mean by a "reason" other than an explanation?

Read through this lens, Williams' juxtaposing of explanation and justification strongly suggests Davidsonian influence. This is corroborated by his final published comments on the argument, which appeal

explicitly to Davidson's concept of a rationalizing explanation (2001: 93).⁴⁵ His initial concern in IER is that, in the case in which an agent drinks petrol because he falsely believes it to be gin, "it is just very odd to say that he has a reason to drink this stuff", despite the fact that "we not only have an explanation of his doing so (a reason why he did it), but we have such an explanation which is of the reason-for-action form" (p. 102). Here he's observing a problem for the Davidsonian analysis, which would have us "ignore the intuition" that saying such a thing is odd. Williams objects that this "looks in the wrong direction, by implying ... that the internal reason conception is *only* concerned with explanation, and not *at all* with the agent's rationality" (p. 102–3, my emphases). The adverbs I've emphasized here suggest the concession to Davidson that the concept of internal reasons is *in part* about explanation. When he turns to address external reasons, he writes, "in considering what an external reason statement might *mean*, we have to remember the dimension of possible explanation" (p. 106, my emphasis).

Williams seeks to accommodate this normative dimension of reasons statements *not* by differentiating between reasons that explain and reasons that justify,⁴⁶ but by distinguishing a normativized kind of explanation: an explanation not of why an action actually was performed but of why it *would* be performed under certain conditions, the conditions of *sound deliberation*.⁴⁷ Introducing this condition, he later writes, "is already enough for the notion to be normative" (1989: 36). We might reconstruct his reasoning as follows. When agents act for

45. Also, in 2001, he writes about "what can rationalize or render intelligible various kinds of action" (p. 94).

46. "It must be a mistake simply to separate explanatory and normative reasons" (1989: 39); "Some writers make a distinction between 'normative' and 'explanatory' reasons, but this does not seem to me to be helpful, because normative and explanatory considerations are closely involved with each other" (2001: 93). For similar scruples, see Darwall 1983; Parfit 1997; Dancy 2000; Finlay 2006.

47. This is supported by Williams' later characterizations of second-personal reasons statements, which he translates as saying, "If I were you, I would ϕ " (1989: 36; 1995: 215).

normative reasons, those are the reasons for which they act. The reasons for which they act just are a certain kind of explanation of their actions. However, sometimes agents act for "reasons" that aren't normative, because they don't justify action, being based on false beliefs. And sometimes agents fail to act for their normative reasons, because they are unaware of them. So normative reasons aren't *simply* explanations of action, because having one neither entails nor is entailed by acting. However, there is a straightforward solution to hand: these cases where there is a lack of entailment between reasons and actions are cases in which the agent mistakenly acts or fails to act due to *false belief* or *ignorance*. The plausible solution, therefore, is that the notion of a normative reason is roughly the notion of an *explanation of an agent's action under the condition of the absence of false belief or ignorance* (i. e., "sound deliberation"). The concept of a normative reason for A to ϕ is the concept of an explanation for why A would have ϕ -ed (or, more precisely, would have been somewhat motivated to ϕ)⁴⁸ if the flaws in his belief-set had not prevented it.⁴⁹ To say that the fact that *this glass contains gin* is a reason for A to drink from it, for example, is to say that it explains why A would drink from it under the condition of sound deliberation. The first premise of Williams' argument is therefore that ' R is a reason for A to ϕ ' means that *R is an explanation of why A would be motivated to ϕ if he deliberated soundly*.⁵⁰

48. Williams often observes the problem that we can have *defeated* reasons, on which we ought not to act. His solution is that these explain why we would have motivation that is outweighed. I shall ignore this complication.

49. Parfit (1997: 110n) suggests that Williams later disowns any such conceptual thesis when he writes, "I certainly did not want to say that [to believe that A has a reason to ϕ] A would have to come to believe *the statement* [that if A deliberated correctly, he would be motivated to ϕ] through deliberation. That would be a very implausible idea" (1995: 188, emphasis in original). But the context suggests a different interpretation. Williams is responding to McDowell's objection that coming to engage in "correct deliberation" might require conversion rather than deliberation. Williams would then be denying he was claiming that beliefs about reasons concern "correct deliberation", *whatever that might mean*, but rather concern something specific he intended by "correct deliberation" (Cf. Skorupski 2006: 79n).

50. "On my view a statement about A 's reasons is partly a statement about A 's

Why analyze normative reasons in terms of these idealizing conditions and not others more favorable to external reasons, like the possession of moral or prudential desires? This choice may look no less question-begging than some versions of the Standard Argument, but Williams offers a justification in his later discussions. He insists (against McDowell's version of this objection) that a statement of reasons for *A* must be distinctively about the actual *A*, and not about some idealized counterpart of *A* (1995: 189–94). But isn't moving to the condition of sound deliberation itself a move to an idealized counterpart of *A*? Williams explains, "The grounds for making this general point about fact and reasoning, as distinct from prudential and moral considerations, are quite simple: any rational deliberative agent has in his [motivational set] a general interest in being factually and rationally correctly informed" (1989: 37; also 2001: 92). I think this should be taken as a claim more distinctly about deliberation than about rationality: any agent, insofar as he deliberates, can be assumed to value the conditions of *sound* deliberation.⁵¹ Williams builds these conditions into the analysis because they are conditions for decision-making that deliberating agents as such can be assumed to value. But why isn't this just as question-begging against external reasons?

The answer lies in the special role and "force" that Williams thinks distinguishes reasons statements from other kinds of normative statements (p. 110; 1989: 39–40; 1995: 191–2). We make claims about "reasons" in the first instance for the purpose of influencing, motivating, or guiding agents. Williams thinks this is also the essential purpose of external reasons claims; in his example Owen Wingrave's family make an external reason claim in order to "urge on him the necessity and importance of his joining the army" (p. 106), and Williams argues that external reasons claims ultimately amount to bluff, or ruses for getting agents to do what the speaker wants (p. 111). But if telling *A* that he

psychology" (1995: 191).

51. Similar claims are made in Smith 1994 and Joyce 2001. Note that Williams allows for rare cases when an agent has a special interest in not being fully informed.

has a reason to ϕ is just to tell him that there is an explanation why he would ϕ under some condition *C*, then we could only expect the reasons claim to influence *A* in favor of ϕ -ing if we thought that he actually values *C* as a condition for making decisions. We would not place an amoralist under any pressure to do the Right Thing, for example, by telling her merely that there is an explanation why she would do the Right Thing if only she cared about morality. To tell her that she has a *reason* to do the Right Thing, however, does seem designed to apply pressure, so it must invoke a condition *C* that even the amoralist is assumed to value. Additionally, we can observe that since only appeals to this condition and not others are sufficient to motivate agents in general, they are thereby uniquely qualified to be explanations (and hence "reasons") for why agents actually act.

Williams' argument depends on a further condition on something's being not merely a reason *why* *A* might ϕ , but also a reason *for* *A* to ϕ . He writes that "what we can correctly ascribe to him in a third-personal internal reasons statement is also what he can ascribe to himself as a result of deliberation" (p. 103), and "internal reasons statements can be discovered in deliberative reasoning" (p. 104). A reason for *A* to ϕ must therefore be an explanation of why *A* would ϕ if he deliberated soundly that *A* himself could discover through deliberation; *i. e.*, it must be an explanation *for* *A* of why he would ϕ . Plausibly, if by 'a reason' we mean *an explanation*, then by 'a reason for *A*' we would mean *an explanation for A*; *i. e.*, something that could be explanatory *to A*. (The expression, '*A* has a reason', which Williams contrasts with 'There is a reason', is also suggestive here of *A*'s *having* an explanation).⁵²

Williams therefore thinks that at least some deliberation takes the form of reasoning to a conclusion about what one has reason to do. Given his conceptual premise, this is to think that to deliberate is sometimes to reason toward an explanation of what one would do if

52. "For it to be the case that he actually has such a reason, however, it seems that the relevance of the unknown fact to his actions has to be fairly close and immediate; otherwise one merely says that *A* would have a reason to ϕ if he knew the fact" (p. 103).

one deliberated soundly. (This couldn't exhaust the possible kinds of deliberation, or else this characterization of deliberation would be vacuous. Fortunately, Williams never suggests that it does).⁵³ This model of deliberation differs strikingly from Humean instrumentalism; Williams writes that "the mere discovery that some course of action is the causal means to an end is not in itself a piece of practical reasoning," whereas "a clear example of practical reasoning is that leading to the conclusion *that one has reason to ϕ because ϕ -ing would be the most convenient, economical, pleasant etc. way of satisfying some element in [one's motivational set]*". If this is right, then Williams' view of deliberation turns out to be surprisingly close to that of "cognitivists" about practical reason like David Velleman and Kieran Setiya, who more recently champion the idea that deliberation aims at acting in a self-comprehensible way.⁵⁴

This model of the conclusion of deliberation predicts a certain kind of view of its premises. The materials for this deliberation would have to take the form of evidence for the explanatory conclusion: evidence that the outcome of sound deliberation would be to ϕ , and why. Indeed, Williams provides exactly such a model of deliberation: the elements in an agent's motivational set ('*S*') appear in the "foreground", as the content of the agent's reasoning.⁵⁵ He writes, "A member of *S*, *D*, will not give *A* a reason for ϕ -ing if ... *A*'s belief in the relevance of ϕ -ing to the satisfaction of *D* is false" (p. 103, my emphasis). This foreground role for elements of the motivational set is apparent in Williams' examples of practical reasoning: "a clear example ... is that leading to the

53. See also Pettit & Smith (2006: 144–5). However, they suggest that Williams' rejection of external reasons is based on the other kinds of deliberation (instrumental deliberation, in their view), whereas I think it is rather this kind of deliberation — the kind he discusses — that is pivotal for the argument.

54. Velleman 1996: 720–4; Setiya 2007: 42. Unlike the "cognitivists", however, Williams does not claim that the beliefs in which practical reasoning concludes are intentions.

55. This suggests an explanation for Williams' talk of "subjective" motivational sets — a seemingly redundant qualifier. (What could an "objective motivational set" be?) Perhaps he means the motivational set that an agent *believes* himself to have. (However, this sits uneasily with the quote from p. 103.)

conclusion that one has reason to ϕ because ϕ -ing would be the most convenient ... way of satisfying some element in *S*"; "thinking how the satisfaction of elements in *S* can be combined"; "where there is some irresolvable conflict among the elements of *S*, considering which one attaches most weight to"; "deciding what would make for an entertaining evening, granted that one wants entertainment". This view of the motivational set as comprising the material for practical thought stands in contrast to the more familiar internalist conception of motives as forces operating in the background.⁵⁶ If Williams were offering some version of the Standard Argument this would be an anomaly, but it is the natural theory of deliberation to embrace on this alternative reading. Just as he maintains, Williams thus does *not* assume an instrumentalist theory of practical reason. He rather emphasizes a model that is evidential.

Some striking evidence for this model of deliberation is in the following passage;

[A]n unknown element in *S*, *D*, will provide a reason for *A* to ϕ only if ϕ -ing is rationally related to *D*; that is to say, roughly, a project to ϕ could be the answer to a *deliberative question formed in part by D*. If *D* is unknown to *A* because it is in the unconscious, it may well not satisfy this condition ... (p. 103, my emphasis)

Here Williams suggests that some element *D* in an agent's motivational set might *not* ground any internal reasons if the agent is unaware of it. This is not because *D*, since it is unconscious, could not be efficacious as a motive; he immediately allows that "of course it may provide the reason why he ϕ 's, that is, may explain or help to explain his ϕ -ing". *D* may be the reason *why A ϕ 's*, but it would not be the reason *for which* he ϕ 's. This passage shows that *contra* some standard readings, Williams is *not* concerned with the mere possibility of explaining an agent's behaviour or of an agent's being motivated. Restricting an

56. Compare Wallace 1999: 223.

agent's reasons by his epistemic access to his motivational set would be peculiar and unmotivated if Williams' guiding thought was simply that reasons must have the capacity to motivate. On the other hand, this epistemic restriction is natural if an agent's reasons are only those considerations that can be explanatory of his actions *to him*.⁵⁷

After sketching this picture of the concept of a reason and the nature of deliberation, Williams turns on pp. 106–9 to the possibility of “external” reasons. His question concerns “what an external reason statement might mean” (p. 106). (He does not ask, How could external reasons motivate?) We must first get clear on what Williams means by calling a reasons statement “external” rather than “internal”. “Internal reasons” are commonly interpreted as reasons for which an agent has a motive that could lead him to act on that reason, and “external reasons” are interpreted as reasons for which an agent lacks any such motive.⁵⁸ This is not what Williams says, however. An internal reason statement is indeed one that is “falsified by the absence” of such a motive (p. 102), but external reasons statements merely “can be true independently” of such a motive (p. 107, my emphasis; also see p. 101; 1989: 35). Internal reasons are thus conceptually dependent on the existence of a relevant motive, while external reasons are not. But this allows that an agent *could* have a motive for acting on an external reason.⁵⁹ Williams is therefore *not* begging the question against external reasons, when he assumes in IER that if something is a reason for *A* then it must be possible for *A* sometimes to be motivated by it.

Williams' case starts (p. 106) with an appeal to the “dimension of possible explanation”; if *R* is a reason for *A* to ϕ , then it could be the

57. Williams' argument is therefore immune to another objection (Millgram 1996: 201–4; Shafer-Landau 2003: 180; Cohon 1993: 271; Johnson 1999; Sobel 2001b), that his case against external reasons is compromised because even some *internal* reasons may be incapable of motivating *A* — reasons that only exist in virtue of *A*'s inability to deliberate soundly (see note 17). Williams would deny that these are genuinely internal reasons, despite their connection to the agent's motivational set: “For it to be the case that he has a reason to ϕ ... the relevance of the unknown fact to his actions has to be fairly close and immediate” (p. 103).

58. For example Brunero 2007: 24.

59. Also see Velleman 1996: 696n; Parfit 1997: 100.

reason for which *A* ϕ 's and hence the explanation of *A*'s ϕ -ing. But “no external reason statement could *by itself* offer an explanation”, Williams observes, since “they can be true independently of the agent's motivations”, and “nothing can explain an agent's ... actions except something that motivates him so to act”. This is not yet an objection to external reasons, however; it merely means that in addition to an external reason statement, “something else is needed ... to explain action, some psychological link”. This psychological link, he is happy to allow, “would seem to be belief”; “*A*'s believing an external reason statement about himself may help to explain his action” (p. 107).

His case against external reasons turns on what it is that *A* comes to believe when he accepts a “specific” reasons claim (p. 107). We can distinguish three different beliefs that *A* might have: (B_1) the *nonnormative* belief that *R* (e.g., that *this glass contains gin*); (B_2) the *specific normative* belief that *R is a reason for A to ϕ* ; and (B_3) the *general normative* belief that *there is a reason for A to ϕ* . Williams has argued, you'll recall, that in order for *R* to be a reason for *A*, *A* has to be able to recognize it as a reason through deliberation. *R* can be a reason for *A* to ϕ , therefore, only if *A* can come to have the specific normative belief B_2 through deliberation. This is a belief that must be able to motivate *A* if it is to be possible that *R* could be the reason for which *A* ϕ 's, as opposed to a mere reason *why A ϕ 's*.

Williams raises no objection to the suggestion that actions can be explained by appeal to beliefs. Indeed he writes, “The claim is ... so plausible, that this agent, with this belief, appears to be one about whom, now, an *internal* reason statement could truly be made” (p. 107). This is not, however, supposed to be in itself a problem for external reasons (“it doesn't follow from this that there is nothing in external reasons statements”, p. 107–8). *A*'s external reason would be *R*, while his internal reason would be *that R is an external reason for him to ϕ* , by virtue of his disposition to be motivated by beliefs of that kind. The significance of this for external reasons statements, Williams tells us, is “that their content is not going to be revealed by considering merely the state of one who believes such a statement, nor how that state explains action” (p. 108). Williams' point here is that it would be a

mistake to think that the possibility of external reasons can be proven just by establishing that external reasons beliefs can motivate.⁶⁰ The external reasons theorist rather needs to show that the agent could come to believe something that would genuinely be a true external reasons statement.

The case comes together on pp. 108–9. For *R* to be a reason for which *A* acts, “coming to believe it must ... involve acquiring a new motivation”. Williams writes that his “basic point” is that the defender of external reasons “must conceive in a special way the connexion between acquiring a motivation and coming to believe the reason statement”. Here he concedes that the motivation might come about in a number of different ways. But the external reasons theorist needs two conditions to be met: (1) “the agent should acquire the motivation because he comes to believe the reasons statement” (p. 108–9; also see 1995: 188n) — otherwise *R* isn’t the reason for which the agent acts; and (2) this must occur “because ... he is considering the matter aright”, *i. e.*, deliberating “correctly”. I suggest Williams here means just that *R* is a reason for *A* only if *A*’s belief that it is a reason for him is true and warranted. It is at this point that Williams makes his peculiarly abrupt claim about what external reasons statements must mean. On the present reading, however, this claim is neither peculiar nor abrupt. Granting his premise about the conceptual content of “a reason for action”, it indeed follows that “the external reasons statement itself will have to be taken as roughly equivalent to, or at least as entailing, the claim that if the agent rationally deliberated, then, whatever motivations he originally had, he would come to be motivated to ϕ ” (p. 109). More precisely, since to believe that *R* is a reason for *A* to ϕ is to believe that *R* is an explanation of why *A* would ϕ if he were to deliberate soundly, what the agent must come correctly to believe is that *R* is an explanation to him, independently of anything in his present motivational set, of why he would be motivated to ϕ if he deliberated soundly.⁶¹

60. Indeed, this is what many “externalists” (in Williams’ sense) seem to have assumed. See Nagel 1970; Darwall 1983. Many commentators make exactly this move as a response to Williams.

61. Admittedly, Williams does not mention explanation here. This is easily

From these requirements, Williams draws his conclusion that “all external reasons statements are false”, because “*ex hypothesi*, there is no motivation for the agent to deliberate from, to reach this new motivation”. Given his evidential model of deliberating to reasons beliefs, what he means is that there are no facts about the agent’s motivational set that warrant his inferring the fact that he has an (external) reason, *i. e.*, that there is an explanation why he would be motivated to ϕ by deliberating soundly, whatever his motivational set. This interpretation takes Williams at his word when he rejects the “sub-Humean” model and professes liberalism about the motivational set. For this argument is strongest if we include dispositions and motivating beliefs in the motivational set. His point is that if an agent cannot discern in his psychology even a disposition to be motivated to ϕ by *R*, then he cannot properly conclude that *R* is an explanation for why he would ϕ if he deliberated soundly, and hence a reason for him to ϕ . Instrumental reasoning from one’s mere dispositions may be an absurdity, but evidential reasoning from them is not.⁶²

Recall that Williams does not deny that an agent could have something in his motivational set that would lead him to be motivated by an external reason; he claims merely that external reasons are independent of any such item. An external reasons statement is therefore

accounted for. He is offering a translation of the general reasons statement (“there is a reason ...”) rather than the specific statement (“*R* is a reason ...”). This would mean there is an explanation why *A* would be motivated to ϕ if he deliberated soundly. (i) Because it omits the explanation, this effectively says merely that *A* would be motivated to ϕ if he deliberated soundly. (ii) Every time Williams offers such a translation, he qualifies it as “roughly equivalent” to or merely entailed by the reasons statement. (iii) This entailment is all he needs for his argument. However, this formulation is infelicitous. An explanation of action on the basis of “whatever motivations he originally had” could still be an internal explanation, if any reasons are grounded in any motivation whatsoever (Schroeder 2007). An external explanation would rather be independent of any motivations he originally had.

62. We can now also appreciate why Williams claims that the indeterminacy of what counts as a rational process and as a member of one’s motivational set favors the internalist over the externalist (p. 110; also 1989: 38 and 2001: 92). The point is that we need not place any formal constraints on what could count as an explanation of an action, except that it has to appeal to the agent’s psychological makeup.

an explanation of an agent's possible action that is independent of any facts about that agent's actual motivational set. But, Williams objects, there can be no such thing. As he later writes, "When the reason is an explanation of his action, then of course it will be, in some form, in his [motivational set], because certainly — and nobody denies this — what he actually does has to be explained by his [motivational set]" (1989: 39).

By their conceptual independence from the agent's desires and dispositions, external reasons statements thus exclude something essential for any evidence that the agent would be motivated through sound deliberation to ϕ .⁶³ Therefore there can be no evidence that would enable *A* to draw the sound inference that *R* would, independently of his motivational set, explain his being motivated to ϕ if he deliberated soundly, since "if [the external reason statement] was true at all, it was true when [*A*] was not motivated" (p. 107). It follows that although external reasons beliefs might be able to motivate, there still cannot be any external reasons, because the condition that this belief must be true and warranted cannot possibly be met.

Williams ends his argument by considering a reply: mightn't an external reason be a *bona fide* explanation in virtue of a disposition to respond to beliefs about reasons? Then, "when he comes to believe that there is reason for him to ϕ , he is motivated to ϕ , even though, before, he neither had a motive to ϕ , nor any motive related to ϕ -ing in one of the ways considered in the account of deliberation" (p. 109). We can now see that Williams' rejoinder is entirely to the point;

[T]his reply ... reapplies the desire and belief model (roughly speaking) of explanation ... but using a desire and belief *the content of which are in question*. What is it that one comes to believe when he comes to believe that there is a reason for him to ϕ , if it is not the proposition, or

63. Heuer (2004) argues against "Humeans" (among whom she includes Williams) that neither desires nor dispositions are themselves significant explanations of actions. But the argument reconstructed here requires merely that identifying a disposition is a necessary condition on explanation.

something that entails the proposition, that if he deliberated rationally, he would be motivated to act appropriately? *We were asking how any true proposition could have that content*; it cannot help, in answering that, to appeal to a supposed desire which is activated by a belief which has that very content. (p. 109–10, my emphasis)

In other words, the problem with the reply is that it is Munchhausen, suggesting that external reasons can lift themselves out of the swamp by their own hair. The fact that the agent would be motivated, *once he had concluded* that *R* is an external reason to ϕ , cannot be any warrant for his drawing that conclusion in the first place. For the agent could with equal warrant have concluded of *anything* that it was an external reason for him to ϕ . Therefore *R* cannot rightly be considered an explanation to him of why he would ϕ if he deliberated soundly.

What exactly is Williams' target, if this Alternative interpretation is correct? There are at least two kinds of so-called "external reasons theorists" who are *not* intended targets: (i) those claiming that agents can have reasons that don't depend upon their actual desires (Williams agrees), and (ii) those claiming that an agent can have reasons that cannot motivate him, even if he comes to recognize them as reasons. (Williams just assumes in IER that this is false as a conceptual matter, and that his opponents concur). So who is the target?⁶⁴ The view that is refuted, if Williams' argument succeeds, is that an agent can correctly recognize something to be a reason for him independently of any psychological information about himself, a view most prominently espoused prior to IER by Thomas Nagel in his *Possibility of Altruism* (1970), and soon after defended by Stephen Darwall in his *Impartial Reason* (1983).^{65, 66}

64. Williams' paper identifies none and provides no bibliography.

65. Indeed Nagel suggests we can learn psychological facts about ourselves through recognizing what our reasons are; see also Korsgaard 1986, Parfit 1997: 102. This is a fundamental contrast.

66. It is worth observing that even many "Humean" accounts of reasons such as Mark Schroeder's (2007) count as "external" on this interpretation.

If this reconstruction is correct, then remarkably not one of the central objections to the argument of IER canvassed in the first section hits its mark. Williams is not conflating “normative” and “explanatory” reasons but (as he tells us) rather analyzing the former notion in terms of the latter notion. He is not begging the question about the nature of rationality but (as he tells us) basing his argument on an independent foundation. However, his starting premise is not therefore unmotivated and (totally) unconvincing; he offers a careful rationale grounded in a view of the conceptual content of a “reason for action” that at the time was widely accepted. He is not assuming an instrumentalist model of rational deliberation but (as he tells us) is offering an alternative, evidential model. He is not dogmatically assuming a faulty neo-Humean theory of motivation; rather (as he tells us) his argument appeals merely to our *dispositions* and maximally inclusive motivational sets. He is not begging the question against his externalist opponents when he assumes that an agent’s reasons are capable of motivating him, because (as he tells us) the “externalists” he has in mind are with him on this point. His argument against their position is not trivial, because (as he tells us) external reasons beliefs *could* motivate agents and thereby play a role in explaining their actions. He does not turn to a separate, “conceptual content” argument to support his case when his main argument fails, plucking a question-begging analysis out of nowhere and obtusely missing the point; rather (as he tells us) his entire argument makes a conceptual point, and the analysis he suggests is entailed by the conceptual premise he defends at the beginning of IER. He does not equivocate over ‘possible motivation’ or ‘motivational set’. Finally, he does not retract any of his original claims or shift his position in his later discussions (as he tells us), and he is not stubbornly clinging to a thoroughly discredited argument.

One might however worry that the Alternative Argument also begs the question against external reasons in its conceptual premise, that to believe *R* is a reason for *A* to ϕ is to believe *R* is a kind of explanation for why *A* would ϕ . As an alleged conceptual truth, it is a stronger claim than any counterpart on an orthodox reading. I believe Williams

thought that the defenders of external reasons had overlooked the question of what we *mean* by a ‘reason for action’ and that they would be hard-pressed to reject his modified Davidsonian analysis. (Hence his challenge to the externalist to say what *else* reasons statements could mean). He seems to concur with Alan Thomas (2002: 137) that the analytic connection between normative and explanatory reasons “can hardly be defended at all”, because “it is true a priori and self-evident to rational reflection”. Most importantly, on this interpretation “external reasons theorists” can (and, Williams assumes, do) agree that external reasons can motivate and therefore explain action. Hence he is especially concerned to address those who mistakenly think that the conceptual link from reasons to explanation is consistent with the existence of external reasons in virtue of the possibility of being motivated by external reasons beliefs. He takes pains to show that this strategy is incoherent and therefore futile.

Although I have argued that this Alternative Argument is not vulnerable to any of the commentators’ main objections, my purpose has not been to defend it, and this conceptual premise is presumably its weakest point. Even granting that the concept of a “reason” just is the concept of an explanation, it doesn’t follow that a normative reason for action is any kind of explanation *of an agent’s acting*. To register my own view, it may be an explanation rather of why *to* act, or why acting so would be in some respect *good* (Finlay 2006). But the correct analysis of ‘a reason for action’ remains a matter of controversy, and the view I am attributing to Williams seems a legitimate contender. As the rest of this argument seems immune at least to the objections that have previously been raised against him, Williams’ skepticism about “external reasons” would deserve a fresh look.

On examination of the textual evidence, and in light of exegetical constraints of charity, consistency, and author’s authority, it therefore seems difficult to resist the conclusion that the Alternative Argument, and not any version of the Standard Argument, is Williams’ *actual* argument. The reader may be incredulous. How could it have happened that such a central article in the ethical literature has been so

universally and badly misunderstood?⁶⁷ I suggest that Williams' choice to introduce his argument by noting the parallels with Hume's views started readers off on the wrong track. He didn't distinguish his argument clearly enough from the familiar instrumentalist argument. And the availability of the correct interpretation, now buried beneath layers of "orthodox" commentary, has receded further as his point of departure, the Davidsonian analysis of the concept of a reason for action, has passed out of philosophical currency. We can blame the obscurity of "internal reasons".⁶⁸

67. This interpretation has been accused of overlooking Williams' endorsement of others' (Standard) readings, particularly McDowell's (1995): "McDowell has set out the main argument, and I shall not repeat it in detail" (1995: 186). In reply, (1) McDowell's formulation of the argument (1995: 68) is vague and largely follows Williams' own, expressed also in terms of interpretation of reasons statements rather than of necessary conditions for reasons. McDowell seems not to recognize the significance this has for Williams; he comments in a footnote that although Williams argues for a "skepticism about whether reasons statements are ever true on the external interpretation", this could just as well be described as skepticism "about whether there are any external reasons", because "nothing turns on whether we are looking for a classification of sorts of reasons for action or merely a classification of sorts of things that can be said" (1995: 83). In his response, Williams accordingly objects that McDowell has not offered a full proposal about the content of external reasons claims (1995: 191). (2) Williams' endorsement is qualified, as he goes on to say, "I do not think that my argument presupposes what McDowell says that it presupposes", and that a "misunderstanding" has arisen (1995: 187). He then explains how his argument "can be best approached", clarifying the *meaning* of reasons claims by appeal to counterfactuals about motivation.

68. This paper was written with the support of a Charles A. Ryskamp research fellowship from the Andrew W. Mellon Foundation and the ACLS, and with the financial support of the Office of the Provost at the University of Southern California. I am grateful for helpful comments from Gunnar Björnsson, Ingmar Persson, Mark Schroeder, Jake Ross, Folke Tersman, some anonymous referees, and audiences at the University of Gothenburg and Uppsala University where I presented a version of the paper. I thank also the editors of *Philosophers' Imprint*.

References

- John Brunero (2007) "McDowell on External Reasons", *European Journal of Philosophy* 16: 22–42.
- Rachel Cohon (1986) "Are External Reasons Impossible?", *Ethics* 96: 545–56.
- _____ (1993) "Internalism about Reasons for Action", *Pacific Philosophical Quarterly* 74: 265–88.
- Stephen Darwall (1983) *Impartial Reason*. Ithaca: Cornell University Press.
- Donald Davidson (1963) "Actions, Reasons, and Causes", reprinted in *Essays on Actions and Events*. Oxford: Clarendon Press, 1980: 3–19.
- Stephen Finlay (2006) "The Reasons that Matter", *Australasian Journal of Philosophy* 84: 1–20.
- _____ (2007) "Responding to Normativity", in Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics* 2. New York: Oxford University Press: 220–39.
- Ulrike Heuer (2004) "Reasons for Action and Desires", *Philosophical Studies* 121: 43–63.
- Martin Hollis (1989) "The Shape of a Life", in J.E.J. Altham & Ross Harrison (eds.), *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press: 170–84.
- Brad Hooker (1987) "Williams' Argument against External Reasons", *Analysis* 47: 42–4.
- Robert Johnson (1999) "Internal Reasons and the Conditional Fallacy", *Philosophical Quarterly* 49: 53–71.
- Richard Joyce (2001) *The Myth of Morality*. Cambridge: Cambridge University Press.
- Christine Korsgaard (1986) "Skepticism about Practical Reason", *The Journal of Philosophy* 83: 5–25.
- John McDowell (1995) "Might There Be External Reasons?", in J.E.J. Altham & Ross Harrison (eds.), *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press: 68–85.
- _____ (2006) "Response to Philip Pettit and Michael Smith", in C.

- MacDonald & G. MacDonald (eds.), *McDowell and his Critics*. Malden MA: Blackwell: 170–79.
- Al Mele (2003) *Motivation and Agency*. New York: Oxford University Press.
- Elijah Millgram (1996) “Williams’ Argument Against External Reasons”, *Noûs* 30: 197–220.
- Thomas Nagel (1970) *The Possibility of Altruism*. Princeton: Princeton University Press (reprint 1978).
- Derek Parfit (1997) “Reason and Motivation”, *Aristotelian Society Supplementary Volume* 71: 99–130.
- Philip Pettit & Michael Smith (2006) “External Reasons”, in C. MacDonald & G. MacDonald (eds.), *McDowell and his Critics*. Malden MA: Blackwell: 142–70.
- John Robertson (1986) “Internalism about Moral Reasons”, *Pacific Philosophical Quarterly* 67: 124–35.
- T. M. Scanlon (1998) *What We Owe to Each Other*. Cambridge MA: Harvard University Press.
- Mark Schroeder (2007) *Slaves of the Passions*. New York: Oxford University Press.
- Kieran Setiya (2004) “Against Internalism”, *Noûs* 38: 266–98.
- _____ (2007) *Reasons without Rationalism*. Princeton: Princeton University Press.
- Russ Shafer-Landau (2003) *Moral Realism: A Defence*. New York: Oxford University Press.
- John Skorupski (2007) “Internal Reasons and the Scope of Blame”, in Alan Thomas (ed.), *Bernard Williams*. Cambridge: Cambridge University Press: 73–103.
- David Sobel (2001a) “Subjective Accounts of Reasons for Action”, *Ethics* 111: 461–92.
- _____ (2001b) “Explanation, Internalism, and Reasons for Action”, *Social Philosophy & Policy* 18: 218–35.
- Michael Smith (1994) *The Moral Problem*. Oxford: Blackwell.
- _____ (1995) “Internal Reasons”, *Philosophy and Phenomenological Research* 60: 109–131.
- Alan Thomas (2002) “Internal Reasons and Contractualist Impartiality”, *Utilitas* 14: 135–54.
- R. Jay Wallace (1990) “How to Argue About Practical Reason”, *Mind* 99: 355–85.
- _____ (1999) “Three Conceptions of Rational Agency”, *Ethical Theory and Moral Practice* 2: 217–42.
- Bernard Williams (1979), “Internal and External Reasons”, reprinted in *Moral Luck*. Cambridge: Cambridge University Press, 1981: 101–13.
- _____ (1985) *Ethics and the Limits of Philosophy*. Cambridge MA: Harvard University Press.
- _____ (1989) “Internal Reasons and the Obscurity of Blame”, reprinted in *Making Sense of Humanity, and other philosophical papers*. Cambridge: Cambridge University Press, 1995: 35–45.
- _____ (1995) “Replies”, in J.E.J. Altham & Ross Harrison (eds.), *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press: 185–224.
- _____ (2001) “Postscript: Some Further Notes on Internal and External Reasons”, in Elijah Millgram (ed.), *Varieties of Practical Reasoning*. Cambridge MA: MIT Press: 91–7.
- J. David Velleman (1996) “The Possibility of Practical Reason”, *Ethics* 106: 694–726.