

THESIS

WILLIAMS ON EXTERNAL REASONS

Submitted By

Marcus Wayne Viney

Department of Philosophy

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, CO

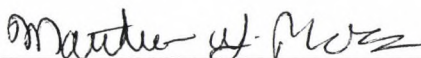
Summer, 2010

COLORADO STATE UNIVERSITY

May 14, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR  
SUPERVISION BY MARCUS VINEY ENTITLED WILLIAMS ON EXTERNAL REASONS BE  
ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF  
ARTS.

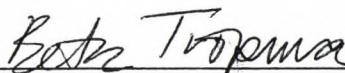
Committee on Graduate Work



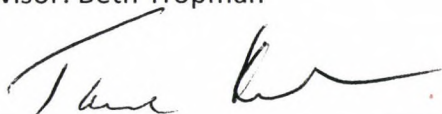
Matthew Rhodes



Michael Losonsky



Advisor: Beth Tropman



Department Head: Jane Kneller

## ABSTRACT OF THESIS

### WILLIAMS ON EXTERNAL REASONS

Bernard Williams has argued for the controversial thesis that there are no external reasons for action. External reasons are desire-independent reasons for action, i.e. reasons whose existence does not depend on the desires of an agent. The thrust of Williams' argument is that an agent's reasons for action must always depend on his or her desires. The overall purpose of this thesis is to clarify and critically examine Williams' argument against external reasons.

In chapter 1 I formalize and explain Williams' argument step-by-step. In addition to this I confront one prominent objection to the argument's validity, which alleges that it contains an equivocation on the term "reason." I argue that this objection fails and that Williams' argument is valid.

In chapter 2 I turn to the soundness of Williams' argument and examine the truth of the premises one by one. In doing this I attempt to uncover important assumptions that underlie Williams' reasoning. I confront several objections to the premises, but I argue that none of them succeed in blocking Williams' conclusion.

In the final chapter I consider the wider issues facing Williams' argument. First I confront three objections which allege that Williams' conclusion has certain

unacceptable consequences. I draw from Williams' work to exonerate his argument on all three counts. Second I discuss two ways the critic might grant the soundness of Williams' argument, but neutralize the impact of his conclusion. While I defend Williams' argument on nearly every point, my primary aim is not to offer a definitive case for the argument. Rather my aim is simply to show that Williams' argument is stronger than some critics might suppose and that it is worthy of further consideration.

Marcus Wayne Viney  
Philosophy Department  
Colorado State University  
Fort Collins, CO 80523  
Summer 2010



## TABLE OF CONTENTS

INTRODUCTION .....	1
CHAPTER 1: WILLIAMS' ARGUMENT AND ITS VALIDITY .....	5
§1. Internal and External Reasons.....	5
§2. Reasons Internalism .....	9
§3. The Significance of Williams' View.....	15
§4. Williams' Argument.....	19
§5. Shafer-Landau's Objection .....	31
§6. Reply to Shafer-Landau .....	34
CHAPTER 2: THE SOUNDNESS OF WILLIAMS' ARGUMENT .....	38
§1. Premise 1.....	39
§2. Reasons and Impossible Actions .....	44
§3. Two Arguments for (R) .....	47
§4. Premise 2.....	53
§5. Cohon's Objection .....	56
§6. Desires and Motivation .....	62
§7. Anti-Humeanism.....	66
§8. Smith's Argument.....	68
CHAPTER 3: THE WIDER ISSUES OF WILLIAMS' ARGUMENT .....	76
§1. The Burden of Explanation.....	77
§2. The Problem of Instrumentalism .....	85
§3. The Problem of Egoism .....	93
§4. Universal Reasons .....	99
§5. Conclusion .....	105

## INTRODUCTION

In "Internal and External Reasons" Bernard Williams argues for the controversial thesis that there are no external reasons for action.<sup>1</sup> Very roughly, according to Williams, external reasons are desire-independent reasons for action, i.e. reasons whose existence does not depend on the desires of an agent. Internal reasons, on the other hand, are desire-dependent reasons for action, i.e. reasons whose existence does depend on the desires of an agent. Accordingly, the thrust of Williams' argument is that all reasons for action are internal, that is, an agent's reasons for action always depend on his or her desires. Williams' conclusion is controversial, because it contradicts a widely-held view that we have at least some reasons for action regardless of our contingent desires.

Many of our reasons for action seem to be internal in Williams' sense. If Marsha desires to bake apple pies for her family, then it seems that she has at least some reason to plant an apple tree, because doing so would be a good way to contribute to the satisfaction of her desire. Marsha's reason may disappear, however, if she loses her desire to bake apple pies. In fact, it would seem rather inappropriate to insist that Marsha has a reason to plant an apple tree even if she has no desire that will be served

---

<sup>1</sup> Bernard Williams, "Internal and External Reasons," reprinted in *Moral Luck* (Cambridge: Cambridge University Press, 1981).



or furthered by doing so. The interesting philosophical question is whether all reasons for action depend on desires in the same way. Williams thinks they do.

One prominent line of opposition to Williams' view grows out of the concern that *moral* reasons for action, for example, seem not to depend on agents' desires. If Marsha happens upon an injured person in need, one might insist that Marsha has a reason to help this person even if doing so will not serve or further any of her desires. In fact, for some, desires will seem to be entirely irrelevant to whether Marsha, or anyone else for that matter, has a reason to help this person. In spite of intuitions like this, Williams thinks that desires are always relevant to whether an agent has a reason for action, moral or otherwise.

The overall purpose of this thesis is to clarify and critically examine Williams' argument for this view. In chapter 1, I will begin by describing Williams' conception of internal and external reasons. From here I will introduce and explain Williams' argument itself, and support my formalization of it against the backdrop of other characterizations of Williams' argument. Finally, I will consider one prominent type of objection to the argument's validity from Russ Shafer-Landau, among others, which contends that the argument contains an equivocation on the term "reason." I will argue that this objection fails and ultimately that Williams' argument is valid.

In chapter 2, I set out to make a case for the soundness of Williams' argument. I will confront several objections to the premises, but I will keep to one premise at a time. In doing this, I will attempt to bring out important assumptions needed to support the premises themselves. I will argue that none of the objections succeed and ultimately

that Williams' argument is sound. As will become clear, however, my aim is not to provide an airtight case for the soundness of Williams' argument. Rather my goal is simply to show that Williams' argument is stronger than some critics might suppose.

In chapter 3, I will confront the wider issues facing Williams' conclusion. First, I will return to the potential consequences of Williams' conclusion for the moral domain. I will draw from Williams' work to show that his conclusion need not be as ruinous as one might assume. Second, I will confront two distinct objections which allege that Williams' conclusion implies instrumentalism about practical reason, and a form of egoism. For the sake of argument, I will grant that these views are problematic, but I argue that neither one is a necessary consequence of Williams' conclusion. Finally, I will consider two plausible ways one might grant the soundness of Williams' argument, but neutralize the impact of his conclusion. Supposing my arguments are on track, I hope to show generally that Williams' argument is by no means easily dismissible and that it deserves further consideration.





## CHAPTER 1: WILLIAMS' ARGUMENT AND ITS VALIDITY

The purpose of this chapter is to describe and develop Williams' argument against external reasons, and ultimately defend its validity. To these ends, it will be necessary to build up Williams' conception of internal and external reasons, with a particular emphasis on Williams' own brand of what I will be calling *reasons internalism*. In addition to this, I will include some general remarks on the philosophical significance of Williams' argument.

### §1. Internal and External Reasons

Williams' conception of internal and external reasons is drawn from ordinary discourse about reasons for action.<sup>2</sup> We often say things like "Marsha had a reason to skip class today," or "Franklin had no reason to break his promise to Tommy." In such cases we are making claims about the reasons for action we believe people have or do not have. And, like any other claims we make about the world, we expect that our reason statements can be true or false. However, with a straightforward claim like "the door in the hallway is open," we can specify with little controversy the conditions under which it is true. We say the proposition "the door in the hallway is open" is true just in

---

<sup>2</sup> The term "reason" is ambiguous; it can refer to a cause, an explanation, a justification, a basis for a conclusion or belief, a process, an activity, a capacity, a faculty, etc. For convenience, I will often use the singular term "reason" to refer to a "reason for action," unless otherwise specified.

case the door in the hallway is in fact open. The truth conditions of reason statements, on the other hand, are a bit trickier.

According to Williams, there are two ways we can understand the truth-conditions of reason statements like “A has a reason to  $\phi$ ” (where ‘A’ is some agent and ‘ $\phi$ ’ is some action).<sup>3</sup> On the “internal interpretation,” the truth of a reason statement is dependent on some element of an agent’s “subjective motivational set,” or *S* for short.<sup>4</sup> For convenience, Williams uses the term “desire” to cover any element of an agent’s *S*, but he insists we not forget that an agent’s *S* may include such things as “dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent.”<sup>5</sup> In essence, the internal interpretation says that the truth of a reason statement implies the presence of a desire on the part of the agent, so that if there is no such desire, the reason statement will be false. Let’s consider an example to make this clear.

A man is accused of lying. He denies this accusation and makes the claim “I have no reason to lie.” Here we have an example of a simple, first-person reason statement that can be true or false. Now, if we are skeptical of the man’s claim, don’t we normally

---

<sup>3</sup> It is unclear what constitutes a single action. If someone decides to go to bed and does so, then this person has acted. But one might say that this person’s action is really composed of several sub-actions, e.g. turning off the lights, walking to the bed, lying down, etc. For our purposes, we can follow Williams and rely on an ordinary understanding of the term “action,” but especially as it denotes *intentional* actions. Thus, intentionally going to bed would be an example of the kind of action we are interested in, while the constituent, apparently mindless motions of this action are not. This is an issue that we will touch on again in chapter 2, section 5.

<sup>4</sup> Williams, *Moral Luck*, 101-102.

<sup>5</sup> I believe Williams goes out of his way to make this qualification because of the egoistic connotations of the term “desire.” For Williams, desires need not be self-serving or egoistic. This is an issue I will return to in chapter 3, section 3. *Ibid.*, 105.



think we can falsify it by pointing out some motive or desire he has to lie? Suppose, for example, we find that the man stands to gain a large sum of money by lying. If this is the case, it certainly does not imply the man did lie, but I think it does imply that the man's original claim is false, that is, that in fact he does have at least some reason to lie. On the other hand, suppose we can find no evidence the man has any motive or desire to lie. In this case, I think it is only reasonable to conclude that the man's original claim is true, that is, there is no reason for him to lie.

It is for cases like these, where there is a direct relationship between an agent's desires and reasons, that the internal interpretation earns its plausibility. I think it is very often that we ascribe or retract reasons for action to ourselves or others on the basis of our desires and motives. However, as Williams recognizes, there are cases in which we intend for a reason statement to be true of someone regardless of his or her desires. Cases such as these form the basis of the "external interpretation" of reason statements. Williams himself characterizes the external interpretation as no more than a mere rejection of the internal interpretation, so, in short, the external interpretation denies that the truth of a reason statement is dependent on an element of an agent's S.<sup>6</sup> Let's consider another example to make this clear.

A man is urged by the police to tell the truth about some matter. Perhaps, for instance, he is the only person with certain information that can save a young girl's life. The man objects and claims, "I have no reason to tell you anything." Here we will stipulate that the man in this case really has no motive or desire to tell the truth. What,

---

<sup>6</sup> Ibid., 101.

then, should we say about the truth of the man's claim? Does he have the reason to talk or not? Falsifying the man's claim in this case is much more difficult, since we cannot simply identify some motivation he has to tell the truth. Now, we would be able to do this, for instance, if the man knew he would go to jail if he didn't tell the truth and he had the desire to stay out of jail. But, by hypothesis, the man has no such desire. Does this mean we should accept the man's claim, and agree that in fact he has no reason to give up the information? I think many people would cringe at the thought of saying the man has no reason to tell the truth here, especially if it is the only way to save a young girl's life. Thus, despite that the man has no relevant motives or desires, the natural thing to say seems to be that the man's original claim is false, that is, that in fact he has a reason to tell the truth. We might even be inclined to add that the man has this reason no matter his desires, because helping the police in a situation like this is simply the right thing to do.

We can see here that the central difference between the internal and the external interpretation of reason statements is the relevance or irrelevance of the elements of an agent's *S*. The internal interpretation holds that for a reason statement to be true of an agent, that agent must have some desire or other, which gives her the reason to act. The external interpretation denies this, and maintains that a reason statement may be true of an agent regardless of his or her desires. Although Williams thinks that both interpretations accurately represent the ways in which we ordinarily understand reason discourse, he rejects that reason statements are ever actually true



on the external interpretation.<sup>7</sup> In other words, Williams thinks, we are systematically mistaken in believing that our external reason statements are ever true.<sup>8</sup> Thus, Williams endorses the internal interpretation as the only correct description of the conditions under which reason statements are true. Before we look at the argument Williams puts forth in favor of this view, it will be helpful to say a bit more about Williams' particular take on the internal interpretation. This will make Williams' argument easier to grasp, but, as we will see, it will also come in handy in dealing with the accusation that his argument is invalid.

To simplify our discussion a little, we may hereafter replace talk of "reason statements on the internal interpretation" with talk simply of internal reasons. If a reason statement is true on the internal interpretation, this just means there exists an internal reason for the agent about which the statement is true. The same goes for reason statements on the external interpretation. To be clear, Williams uses the term "internal" to refer to the internal or necessary connection a reason has to an agent's *S*, and "external" to refer to the lack of such a connection.

## **§2. Reasons Internalism**

Williams' view is commonly referred to as *internalism about reasons*, or more simply *reasons internalism*.<sup>9</sup> Essentially, reasons internalism is the view that all reasons

---

<sup>7</sup> Ibid., 101.

<sup>8</sup> As we will see later, this claim is not left without explanation. See chapter 3, section 1.

<sup>9</sup> In contemporary metaethics, the term "internalism" can refer to a number of different positions. For a helpful disambiguation of this term, see: John Robertson, "Internalism, Practical Reason,



for action are internal, and thus no reasons for action are external. But there are several ways one could spell out this view. The most we have said so far about Williams' particular brand of reasons internalism is that a reason statement is falsified in the absence of an element of an agent's *S*. In other words, if an agent doesn't have a relevant desire, that agent will not thereby have an internal reason to act. This means that Williams thinks having a desire is a necessary condition for having a reason to act. But this may make us wonder whether having a desire is not also a sufficient condition for having a reason to act. That is, if an agent has a relevant desire, will this always thereby give that agent an internal reason to act? According to Williams, this is what the most elementary version of reasons internalism would state. On what we might call *basic reasons internalism* having a desire is both a necessary and sufficient condition for having a reason to act.<sup>10</sup> This view may be characterized in the following manner:

**Basic reasons internalism:** A has a reason to  $\phi$  if and only if A has a desire that will be satisfied by her  $\phi$ -ing.<sup>11</sup>

We can see quickly, however, that for Williams this version of reasons internalism is too simplistic. First of all, Williams prefers the expression "a desire that will be served or furthered by her  $\phi$ -ing" to "a desire that will be satisfied."<sup>12</sup> Despite that the latter expression is more commonly associated with reasons internalism,

---

and Motivation" in *Varieties of Practical Reasoning*, ed. Elijah Millgrim (Cambridge, Massachusetts: MIT Press, 2001).

<sup>10</sup> Williams refers to this as the "sub-Humean" model of reasons, because it represents an overly simplistic expression of David Hume's position on practical reason. I find it less distracting to refer to this view and its variations as "reasons internalism." Williams, *Moral Luck*, 102.

<sup>11</sup> Ibid., 101.

<sup>12</sup> Ibid.

Williams himself says nothing about the rationale for this preference. Nonetheless, I think a simple example can show what he might have had in mind. Say Marsha desires to become a doctor and knows that she needs to study for her upcoming exam. We may stipulate that this exam is only one of many that she will need to pass in order to become a doctor. In this case, Marsha's original desire will not be satisfied by her studying for this upcoming exam, because there are other actions she will need to take to become a doctor. Nevertheless, I think we would want to say that her reason to study for the exam is ultimately grounded by her original desire. Thus, a desire need not be fully satisfied by one action in order for it to ground a reason for taking that action; according to Williams, the desire need only be served or furthered.

In addition to this consideration, Williams thinks that basic reasons internalism is too simplistic because we do not want to say that having a desire is always enough for an agent to have a reason to act. Williams offers the following case. A man desires a gin and tonic and believes the bottle in front of him contains gin, when in fact it is petrol.<sup>13</sup> Williams' question is this: should we say that the man has a reason to mix the contents of the bottle with tonic and drink it? That is, given that the man has a desire, does he thereby have a reason to act? In a more normal situation, where the bottle did contain gin, I think we might agree that the man does have such a reason. But does the man retain this reason even when the bottle contains petrol? Initially, Williams has this to say about the case:

---

<sup>13</sup> Ibid., 102.



On the one hand, it is just very odd to say that he has a reason to drink this stuff, and natural to say that he has no reason to drink it, although he thinks he has. On the other hand, if he does drink it, 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.<sup>14</sup>

Williams' idea here is that, although we could explain the man's action if he does unknowingly drink the petrol (because we could cite his false belief that the bottle contained gin and his desire for a gin and tonic), we should not say that he in fact had a reason to drink the liquid. This is because, for Williams, reasons internalism ought to be concerned with an agent's rationality, and not merely with explanation.<sup>15</sup> Put differently, Williams does not think a general account of reasons for action should generate reasons where we intuitively do not think there are any, such as the man's reason to drink the petrol. Williams argues that reasons internalism can comply with this constraint by embracing the notion of a "sound deliberative route." For Williams, reasons internalism need not admit that the man in the gin/petrol case has a reason to drink the liquid, because "what we can correctly ascribe to him in a third-personal internal reason statement is also what he can ascribe to himself as a result of [sound] deliberation."<sup>16</sup> In other words, according to Williams, since from our perspective we can see that if the man drinks the contents of the bottle, he will become very ill, it seems reasonable to say that if the man were fully aware of the situation, and if he were to deliberate correctly, he himself would come to see that mixing the liquid with tonic

---

<sup>14</sup> Ibid., 102.

<sup>15</sup> Ibid., 102-103.

<sup>16</sup> Ibid., 103.

would not satisfy his original desire for a gin and tonic (and furthermore that it would likely hinder other more important desires of his, e.g. the desire not to drink petrol, or the desire to live).

Williams proposes two additions to reasons internalism to account for cases like these. First, having a desire is no longer a sufficient condition for having a reason to  $\phi$ , only a necessary one.<sup>17</sup> Second, Williams adds in a rather convoluted fashion, “[an element] of  $S$ ,  $D$ , will not give  $A$  a reason for  $\phi$ -ing if either the existence of  $D$  is dependent on a false belief, or  $A$ ’s belief in the relevance of  $\phi$ -ing to the satisfaction of  $D$  is false.”<sup>18</sup> Stated more simply, for an agent to have a reason to act, two conditions must hold: (1) an agent’s relevant beliefs must all be true, and (2) the agent must be reasoning correctly.<sup>19</sup> Sometime after “Internal and External Reasons,” Williams coined the expression “sound deliberative route” to cover both of these conditions.<sup>20</sup> According to Williams, having a sound deliberative route simply means that  $A$ ’s

---

<sup>17</sup> Although I am dropping the sufficient condition from Williams’ reasons internalism here, it should be noted that he says in passing, “I actually think that it provides a sufficient condition as well...” However, an exploration and defense of Williams’ claim would lead us too far astray from our present purposes. Bernard Williams, “Internal Reasons and the Obscurity of Blame” in *Making Sense of Humanity* (Cambridge: Cambridge University Press, 1995), 35.

<sup>18</sup> Williams, *Moral Luck*, 103.

<sup>19</sup> Presumably Williams has in mind various formal and informal fallacies an agent might commit while deliberating about what to do in a given situation. A good example might be a man who believes he has a reason to gamble over and over again simply because he does not appreciate the gambler’s fallacy.

<sup>20</sup> Bernard Williams, “Postscript: Some Further Notes on Internal and External Reasons,” in *Varieties of Practical Reasoning*, ed. Elijah Millgrim (Cambridge, Massachusetts: MIT Press, 2001), 91.



deliberation about what to do in a given case could begin from true premises and progress soundly to the practical conclusion to  $\phi$ .<sup>21</sup>

Two significant consequences follow from the addition of sound deliberative routes to reasons internalism. First, on Williams' view, it is possible for an agent to believe falsely that he has an internal reason to  $\phi$ , as with the unlucky gin and tonic drinker.<sup>22</sup> This is because there is no sound deliberative route from the man's  $S$  to his drinking the liquid, though he thinks there is. Second, for Williams, it is possible for an agent to be unaware that he has an actual internal reason to  $\phi$ .<sup>23</sup> For example, unbeknownst to Franklin, he has a reason to see the doctor, because he will die if he does not, and he desires to live. Here there is a sound deliberative route from Franklin's  $S$  to his seeing the doctor, but unfortunately he is unaware of it.

One last amendment remains. According to Williams, it is important that we not think of an agent's  $S$  as statically given. This is because, Williams thinks, "the process of deliberation can have all sorts of effect on [an agent's]  $S$ ."<sup>24</sup> For example, an agent could gain a new desire and lose another simply by an exercise of her imagination. Perhaps, for instance, she was not aware of what was really possible in a given situation. In such a case, that agent may acquire a new internal reason to  $\phi$ , or abandon an old one. To capture the variability of such cases, Williams proposes that an agent may have

---

<sup>21</sup> Here "soundly" is meant to cover inductive as well as deductive inferences.

<sup>22</sup> Williams, *Moral Luck*, 103.

<sup>23</sup> *Ibid.*, 103.

<sup>24</sup> *Ibid.*, 105.



reason to  $\phi$  even if there is no desire in her existing  $S$  (that will be served or furthered by her  $\phi$ -ing), so long as there is some path of sound deliberation that could bring her to such a desire.<sup>25</sup>

To display all of these refinements, but especially as they contrast with basic reasons internalism, we may lay down Williams' reasons internalism in the following manner. I prefer one of the last formulations Williams offered, so I will follow suit:

**Williams' reasons internalism:** "A has a reason to  $\phi$  only if... [A] could reach the conclusion to  $\phi$  by a sound deliberative route from the motivations... [A] already has."<sup>26</sup>

Thus far, we have built up Williams' view in some detail, but we have yet to see an argument for why it is true, or why anyone might care whether it is true. Before we turn to Williams' argument, I want to say a few words about why Williams' view is supposed to be interesting, including a brief description of the view that is alleged to be ruled out by Williams' argument.

### §3. The Significance of Williams' View

Williams' view draws attention, good and bad, because of its perceived consequences for the moral domain. There is a common view, both inside and outside of philosophy, which treats moral reasons for action as external in Williams' sense. That is, according to this view, moral reasons for action may apply to agents regardless of their particular desires and motivations. Given that Williams' view states that all reasons for action are internal, it follows that, if Williams is right, this common view

---

<sup>25</sup> Ibid., 104-105.

<sup>26</sup> Williams, *Making Sense of Humanity*, 35.

must somehow be mistaken. For simplicity, I will refer to this view as *reasons externalism*, and define it against the backdrop of our discussion of Williams' view:<sup>27</sup>

**Reasons externalism:** Possibly (A has a reason to  $\phi$  and there is no sound deliberative route from A's actual motivations to A's conclusion to  $\phi$ ).

The most important element in reasons externalism, for our purposes, is that having a desire, in one way or another, is *not* a necessary condition for having a reason to act. In other words, according to reasons externalism, it is possible for an agent have a reason to a  $\phi$ , even if he is not, or could not be, motivated to  $\phi$ . Let's borrow an example from Williams to see why someone might be drawn to this view:

Suppose, for instance, I think someone (I use 'ought' in an unspecific way here) ought to be nicer to his wife. I say, 'You have a reason to be nicer to her.' He says, 'What reason?' I say, 'Because she is your wife.' He says – and he is a very hard case – 'I don't care. Don't you understand? I really do not care.' I try various things on him, and try to involve him in this business; and I find that he really is hard case: there is *nothing* in his motivational set that gives him a reason to be nicer to his wife as things are.<sup>28</sup>

There are two important ideas intertwined in this example. First, we the readers are meant to believe that the husband has a genuine reason to treat his wife better.<sup>29</sup> Second, it is stipulated, there is no sound deliberative route from the husband's *S* to the conclusion to treat his wife better. The problem then, of course, is that if we accept reasons internalism, we must either admit that the man in fact has no such reason,

---

<sup>27</sup> It should be noted that, like reasons internalism, there are a variety of ways one could specify reasons externalism. I choose this characterization because it captures the most general claim I think any reasons externalist would accept. As a side note, this view is often referred to as "externalism," but once again I find this confusing, since there are other views with the same name.

<sup>28</sup> Williams, *Making Sense of Humanity*, 39.

<sup>29</sup> It is important to note that this is not the same as saying "there is a reason for treating one's wife better." For Williams, the claim of interest is more specific: "there is a reason for *this man* to treat his wife better."



which seems to be an unpalatable option, or weaken reasons internalism to make room for the intuition that he has such a reason. But, if we do the latter, we may lose what is really distinctive about reasons internalism in the first place. An essential piece of reasons internalism, as I understand it, is that desires, one way or another, are required for having a reason to act. Accordingly, if reasons internalism allows for the existence of reasons in the absence of desires, then there is nothing particularly interesting about the claim of reasons internalism.

In contrast, this example poses no particular problem for reasons externalism. In fact, reasons externalism seems to accord well with our ordinary intuitions about this moral case and others. We typically do not think, for example, that an agent can excuse himself from moral reasons for action simply because he has no desires that will be served or furthered by acting in this way. Accordingly, with regard to our example from above, despite that there is no sound deliberative route from the husband's *S* to the conclusion to treat his wife better, I think the common intuition would be that the man nevertheless has a reason to do so. And, presumably, the reasons internalist simply cannot account for this intuition. John McDowell, a critic of Williams' view, states the general problem faced by reasons internalism nicely: "The implication of Williams' skepticism is that ethical reasons are reasons only for those for whom they are internal reasons: only for those who have motivations to which ethical considerations speak, or can be made to speak."<sup>30</sup> I think McDowell's basic point here is that, if we accept

---

<sup>30</sup> John McDowell, "Might there be external reasons?" in *World, mind, and ethics: Essays on the ethical philosophy of Bernard Williams*, ed. J.E.J Altham and Ross Harrison (Cambridge: Cambridge University Press, 1995), 68.

Williams' reasons internalism, we must be willing to admit that it would be, in principle, possible for an agent to forgo completely any particular moral reason for action. If an agent genuinely is not, or could not be, motivated to act morally, then, according to Williams, that agent truly has no reason to act morally. Having a desire is always required for having a reason, for Williams, so without some desire or other, there can be no reason – *even if* the reason is moral. For some, giving up this common view of moral reasons would be an unbearable consequence of accepting Williams' view.<sup>31</sup> I suspect this is why many people are, at least initially, magnetized to reasons externalism.

While my goal is ultimately to support Williams' view, I believe it is only reasonable to concede that moral examples, such as the insensitive husband, present a *prima facie* case against Williams' reasons internalism. Do we really want to say that the man has no reason to treat his wife better? Furthermore, do we really want to admit, in general, that moral reasons for action are no more binding than any other kind of reason for action? The intuitive answer, for many, is of course "no." I submit here, however, that if there is a major fault in Williams' reasons internalism, it must be possible to find some flaw in the argument he puts forth in support of it, i.e. it must be possible to show that his argument is either invalid or unsound. And, if no such flaw can be successfully identified – in spite of our troubling intuitions – I think we should agree that Williams' argument wins the day.

I will turn now to Williams' argument, but in way of preview, following this discussion I will confront a prominent objection to Williams' argument, which contends

---

<sup>31</sup> I will treat this issue in more detail in chapter 3, section 1.



that it is invalid. The particular version of the objection I want to examine comes from Shafer-Landau. In essence, Shafer-Landau accuses Williams of committing the fallacy of equivocation due to an alleged confusion over two different senses of the term “reason.” To be sure, Shafer-Landau is not the only person to raise the worry he does, but I choose to confront his version, because I find it to be the strongest and most clearly stated. It is crucial to silence this objection first, because if Williams’ argument was invalid, any arguments over the truth of the premises would be irrelevant.

#### **§4. Williams’ Argument**

Many formulations of Williams’ argument have been offered. The one I present here is my own, but immediately following my discussion I will examine a few other versions to help clarify, and partially justify, my particular take on Williams’ line of thought. Let’s begin with a snapshot of the argument, followed by more a detailed discussion of each step. Williams denies the existence of external reasons on the grounds that they could not possibly explain an agent’s intentional actions. Given that people act for reasons, it follows that such reasons must figure in correct explanations of their actions. How could we explain someone’s intentional action correctly unless we take account of her specific reason for doing so? Yet, according to Williams, the only thing that could explain an agent’s intentional actions are things that could motivate that agent to act. This means that external reasons by themselves could not be the explanation of an agent’s actions, because external reasons are supposed to bear no relation to an agent’s motivations whatsoever.



We may formalize Williams' argument as follows (For all  $R$ ,  $A$ , and  $\phi$ ):<sup>32</sup>

- (1)  $R$  is a reason for  $A$  to  $\phi$  only if  $R$  could explain  $A$ 's  $\phi$ -ing.
- (2)  $R$  could explain  $A$ 's  $\phi$ -ing only if  $R$  is related to  $A$ 's subjective motivational set.
- (3) External reasons are not related to  $A$ 's subjective motivational set.
- (4)  $R$  is not an external reason for  $A$  to  $\phi$ , i.e. there are no external reasons for  $A$  to  $\phi$ .

The central insight of Williams' argument is that all reasons for action must have an "explanatory dimension."<sup>33</sup> What Williams means by this is that all reasons for action must be able to explain the actions performed for those reasons. Williams says, "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."<sup>34</sup> Suppose, for example, that Marsha has a reason to plant an apple tree, because she wants to bake apple pies for her family. In this case, according to Williams, it must be possible that Marsha should plant an apple tree for this reason, and if she does, her reason will figure in a correct explanation of her action. The idea behind this is simple. Given that Marsha's action was intentional, we could not fully capture why she acted unless we reference her specific reason doing so. That is, unless we understand the

---

<sup>32</sup> To be clear, Williams himself never formalizes his argument in this manner. However, I believe it is important to make the form of his argument conspicuous, because it is easy to confuse or collapse the distinct steps of his argument.

<sup>33</sup> Williams, *Moral Luck*, 102.

<sup>34</sup> *Ibid.*, 102.

particular consideration for which she acted, we will miss the most crucial element in the story behind her action. Thus, according to Williams, any reason for action must be able to explain the action performed for that reason.

To be clear, Williams' point here is not that all reasons for action must actually explain the actions performed for those reasons, only that they have the potential to do so. For, it could have been the case that, although Marsha had a reason to plant the apple tree, she never acted for this reason. Perhaps, for instance, Marsha found out it would be easier to buy apple pies from the store than to make them from scratch, and so decided against planting the apple tree. In this case, Marsha's reason for planting the apple tree would not explain her action, because she never acted for it, but, according to Williams, it *could* have explained her action, if she had acted for it, or will act for it in the future.

It is important to be clear about the sense of possibility at work in premise 1. By "could explain" Williams surely does not mean it must be *logically* possible for a reason to explain an agent's action, because this would let too many considerations count as reasons – so long as there is no contradiction involved, presumably any consideration could explain anyone's action. Williams wants to be more restrictive than this, and it seems to me that he must have at least two types of constraint in mind for the sense of possibility expressed in premise 1, namely, nomological and historical necessity. Nomological necessity refers to the constraints imposed by the laws of nature in the actual world, and historical necessity refers to the constraints imposed by the way in which history has unfolded for humans in the actual world. Thus, while it might be



*logically* possible for some consideration, e.g. that Shakespeare's plays are beautiful, to explain Aristotle's reading of Shakespeare's plays, it is not *historically* possible for this consideration to do so, because Aristotle lived and died 1,800 years before Shakespeare wrote his plays. According to premise 1, therefore, the fact that Shakespeare's plays are beautiful *could not* be a reason for Aristotle to read these plays, because this fact *could not* explain his doing so.<sup>35</sup>

The next step of Williams' argument is that reasons for action, in order to have an explanatory dimension, must be related to an agent's subjective motivational set (again, *S* for short). For Williams, any reason that is unrelated to an agent's *S* could not possibly explain any action performed by that agent, since "nothing can explain an agent's (intentional) actions except something that motivates him so to act."<sup>36</sup> It seems that Williams has in mind here some version of the belief-desire theory of action explanation (BDT for short), which holds that both beliefs and desires are necessary to explain intentional actions. Imagine, for example, that Franklin and Tommy are exploring the mountains together and happen upon an irritated rattlesnake. After a moment's thought, Franklin slowly steps away to safety, but Tommy does not. Suppose we want to explain Franklin's action, so we ask him: "Why did you step away just then?" and he explains, "Because I thought the rattlesnake would bite me." This explanation should satisfy our curiosity – that is, until we come to find out that Tommy had the exact same thought, but didn't step away from the snake. In light of this, we must now

---

<sup>35</sup> We will return to reasons and possibility in chapter 2, section 2.

<sup>36</sup> *Ibid.*, 107.



include some further element besides Franklin's belief to explain his stepping away. According to the BDT, this further element must be something motivational, like a desire.<sup>37</sup> Indeed we can explain Franklin's action, and Tommy's inaction, by discovering that while Franklin desires not to be bitten, Tommy does (perhaps, for instance, Tommy is a masochist and always wanted to know what it felt like to be bitten by a rattlesnake).<sup>38</sup>

What this example aims to show is that any consideration, e.g. that a rattlesnake is ready to strike, cannot by itself explain anyone's action, because two people can take into account the exact same consideration when deliberating, but act differently. Thus, according to Williams, in order to explain someone's intentional action fully, we must reference some element of this person's *S* which is related to the consideration at hand.

The last step of Williams' argument is a definitional point about external reasons. For Williams, an external reason is a reason for action which applies to an agent regardless of the elements of this agent's *S*. But, if an external reason bears no relation to an agent's *S*, then, according to the BDT, citing such a reason could not possibly explain any action performed by this agent. Yet, if this is the case, then, according to

---

<sup>37</sup> At this point I am merely introducing these ideas; I will treat them critically in chapter 2, section 6.

<sup>38</sup> One might question why this further element must be a desire. Perhaps Tommy *believes* something Franklin does not, e.g. rattlesnakes will not strike someone wearing red, and this is what explains their difference in action. According to the BDT, this will not work, because we can easily change the situation so that beliefs alone could not possibly explain the difference. Imagine that three days ago Franklin happened upon a rattlesnake and stepped away, and none of his beliefs have changed, but today, when he happened upon a rattlesnake again, he did not step away. In this case, we must include something besides Franklin's beliefs, and according to the BDT, the best candidate for this is a motivational state like desires. Perhaps three days ago Franklin desired not to be bitten, but today he is depressed, and does not care what happens to him.

premise 1, such a reason could be no reason at all. Put differently, there can be no such things as external reasons; only internal reasons can explain actions, because only internal reasons are related to an agent's *S*. Let's consider an example to make this step of the argument clearer.

Williams invites us to consider the case of Owen Wingrave, a man who was born into a family with a longstanding tradition of military service.<sup>39</sup> Owen's father urges upon him the importance of joining the army, because "all of his male ancestors were soldiers, and family pride requires him to do the same."<sup>40</sup> However, we are also told that Owen "...has no motivation to join the army at all, and all his desires lead in another direction: he hates everything about military life and what it means."<sup>41</sup> Despite this, Owen's father insists that "*there... [is] a reason for Owen to join the army.*"<sup>42</sup> Given Owen's situation, we have no choice but to interpret the father's claim *externally*, that is to say, that there is an external reason for Owen to join the army.

For the sake of argument, let's suppose that what the father says is true, and that Owen does in fact have an external reason to join the army, where Owen's reason is this: "[Owen's] family has a tradition of military honor."<sup>43</sup> In considering what this might mean, Williams says, we must keep in mind the "dimension of possible

---

<sup>39</sup> Williams, *Moral Luck*, 106-109.

<sup>40</sup> *Ibid.*, 106.

<sup>41</sup> *Ibid.*

<sup>42</sup> *Ibid.*

<sup>43</sup> *Ibid.*, 107.



explanation" (from premise 1), a requirement that "applies to any reason for action."<sup>44</sup> Accordingly, if there is a reason for Owen to join the army, this reason must have the potential to explain Owen's doing so. So, could Owen's external reason ever explain Owen's joining the army? Initially, Williams thinks not, because, "Even if it were true...that there was [an external] reason for Owen to join the army, that fact by itself would never explain anything that Owen did, not even his joining the army. For if it was true at all, it was true when Owen was not motivated to join the army."<sup>45</sup> In other words, according to Williams, the bare consideration that Owen's family has a tradition of military honor could not possibly serve as an explanation of Owen's joining the army, because this would have been true, if it was at all, even before Owen joined the army. So, supposing Owen did join the army on some particular occasion, and we requested an explanation of his doing so, and someone replied "because his family has a tradition of military honor," we could rightfully respond, "but that was true before he joined; what explains the fact that Owen joined the army only recently?" Since something else is required here, Williams says "no external reason...could *by itself* offer an explanation of anyone's action."<sup>46</sup> By premise 1, for Williams, this implies that external reasons are impossible.<sup>47</sup>

---

<sup>44</sup> Ibid., 106.

<sup>45</sup> Ibid., 106-107.

<sup>46</sup> Emphasis in original. Ibid., 106.

<sup>47</sup> Strictly speaking, premise 1 says nothing about reasons having to explain "by themselves." I will confront this apparent slippage shortly.



But this need not be the end of the story. Perhaps, as Williams suggests, all we need is a “psychological link” between Owen and his external reason to explain his joining the army, like a belief.<sup>48</sup> In other words, perhaps we could explain Owen’s joining the army by citing Owen’s recently acquired belief that considerations of family honor constitute a reason for him to join the army. One might now explain, “Before he joined the army, Owen had no such belief, but once he acquired this belief, he joined right away. So, this is what explains why he did so only recently.” If something like this can be the case, then perhaps external reasons can explain intentional actions after all.

This approach seems plausible at first, but for Williams it will not work. This is because implicit in Williams’ acceptance of the BDT (from premise 2) is a commitment to a version of the Humean theory of motivation (HTM for short), which states that beliefs are not sufficient, and desires are necessary, for motivation. According to the HTM, a belief by itself cannot motivate someone to act, because beliefs are merely *cognitive* states which aim to represent the world. A desire, on the other hand, can motivate someone to act, according to the HTM, because desires are *conative* states which, unlike beliefs, aim to alter the world. Following Michael Smith, the difference between cognitive and conative states, exemplified by beliefs and desires, is commonly treated as a difference between “directions of fit.”<sup>49</sup> When I believe, for example, that there is five dollars in my pocket, and I find out there is no money in my pocket after all, it is my belief that must change. In contrast, when I desire to have five dollars in my

---

<sup>48</sup> Ibid., 107.

<sup>49</sup> Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994), 111-125.

pocket, and I find out there is no money in my pocket after all, it is the world that must change. For this reason, beliefs are said to have a “mind-to-world-fit,” whereas desires are said to have a “world-to-mind-fit.” Since desires tend to persist where a belief would not, desires are taken to be the only thing that can motivate someone to act or to bring about some new state of affairs in the world, e.g. making it true that I have five dollars in my pocket.

To reconnect with Owen’s case, we can see more clearly why Owen’s belief that he has an external reason, alone, cannot explain his joining the army. According to the HTM, beliefs by themselves cannot motivate people to act, so Owen’s new belief cannot fully explain his action, because it is possible for someone else to have the exact same belief, but not be motivated to join the army. To explain why Owen joined the army, we must, in addition to his belief, include some motivational element like a desire. Yet, according to Williams, if we must do this, then any putative external reason collapses into an internal one; a reason with a necessary relation to a desire simply *is* an internal reason. With regard to Owen’s case, Williams thinks, if we must include a desire in our explanation of his joining the army, then Owen “appears to be one about whom, now, an *internal* reason statement could truly be made: he is one with an appropriate motivation in his *S*.”<sup>50</sup> Thus, our original supposition that Owen had an external reason to join the army must be false, because only an internal reason could explain his intentional actions. By premise 1, this means that all reasons for action are internal, and therefore, according to Williams, there is no such thing as an external reason.

---

<sup>50</sup> Williams, *Moral Luck*, 107.



This ends my particular take on Williams' argument against external reasons. Before we consider the first objection, I want to present a few other formulations of the argument to help solidify our understanding of it. This should also show that my interpretation of Williams' argument accords well with the general understanding of the argument as it exists in the literature today. All four of the following passages come from critics of Williams' argument.

The first, offered by Shafer-Landau, is perhaps the closest to my reading:

...our reasons are linked with the possibility of explaining action. Reasons figure in action explanation, but if there were reasons that were unrelated to our subjective motivational sets, then they could not possibly explain our actions. Yet it must be possible to explain our actions by citing our reasons for performing them. Therefore reasons must be related to our subjective motivational sets...<sup>51</sup>

The only problem with this characterization of Williams' argument is that it lacks explicit reference to external reasons. We the readers are left to infer from this that external reasons are impossible.

The next version, from Christine Korsgaard, is an improvement in this manner:

Since an external-reason claim does not imply the existence of a motive, it cannot be used to explain anyone's action: that is, we cannot say that the person *P* did the action *A* because of reason *R*; for *R* does not provide *P* with a motive for doing *A*, and *that* is what we need to explain *P*'s doing *A*: a motive.<sup>52</sup>

It seems to me, however, that Korsgaard's expression of the argument is slightly more cumbersome than need be. Rather than using the roundabout language of "external-reason claims," I prefer to stick with the more direct language of "external reasons." In

---

<sup>51</sup> Russ Shafer-Landau, *Moral Realism* (Oxford: Oxford University Press, 2003), 178.

<sup>52</sup> Christine Korsgaard, "Skepticism about Practical Reason" *The Journal of Philosophy*, 83, no. 1 (January 1986), 11.



the end, it makes no difference to Williams' argument one way or the other, so we might as well go with the clearest route possible. The next version comes from Ulrike Heuer, which is perhaps the most straightforward formulation of the argument available.

Heuer's presentation also stands as one of the only attempts to formalize Williams' argument:

- (1) A practical reason...must be capable of explaining an action that was performed for that reason...
- (2) Actions can only be explained by motivationally relevant attitudes, i.e. desires...
- (3) All practical reasons must be based on desires...<sup>53</sup>

This formulation of Williams' argument should be appreciated for its simplicity. In my opinion, however, the strength of Heuer's version is also its weakness. The precise link between the premises is somewhat sloppy, and the conclusion as it stands is unclear – what exactly does it mean for reasons to be “based on” desires? Furthermore, like Shafer-Landau's formulation, Heuer's passes over any specific points about external reasons. Against the background of this formulation and the other two, the fourth, authored by Elijah Millgram, is the one I prefer the most. This is because it not only avoids the minor faults of the others, but it also incorporates an important aspect of the argument that can easily go unnoticed:

...reasons must be able to explain actions, but to explain actions, one must appeal to motivations; and an explanation that adduces motivations is an internal, and not an external, reasons explanation.<sup>54</sup>

---

<sup>53</sup> Ulrike Heuer, “Reasons and Actions and Desires” *Philosophical Studies*, 121, no. 1 (October 2004), 46.

<sup>54</sup> Elijah Millgram, “Williams' Argument Against External Reasons,” *Noûs* 30, no. 2 (1996), 197.

Like the others, Millgram notes the basic move from reasons to explanations and from explanations to motivations to display the necessary connection between reasons and motives. But unlike the others, Millgram brings out an important point Williams makes, which we saw in our discussion of Owen Wingrave. Recall that, according to Williams, any putative external reason collapses into an internal one, because any reason that can explain an agent's intentional action is one that has a necessary relation to that agent's *S*. Now, one might think that because the definition of an external reason is one that an agent has *regardless* of her desires, an agent could have an external reason which simply happens to align with her desires, and it would therefore have no problem in explaining an action performed by this agent. Perhaps, for example, it was true all along that Owen had an external reason to join the army, and this reason could explain his doing so, because it just so happened that Owen also had a desire to join the army. Doesn't this show that external reasons are possible after all?

Williams would reject that this kind of response can save external reasons. External reasons are reasons that an agent has regardless of her desires. This means that an external reason is either related to an agent's *S* or it is not. If an external reason is not related to an agent's *S*, then, as we have seen, this reason could not explain any action performed by this agent. Of course, in this case, according to Williams, the external reason is no reason at all. Now, on the other hand, if an external reason *is* related to an agent's *S*, because the agent happens to have a desire which lines up with this reason, then, it seems, this reason could explain an action performed by this agent. In this case, however, Williams would agree with Millgram that "an explanation that



adduces motivations is an internal, and not an external, reasons explanation.” In other words, if it is necessary to cite a relevant motivation in order to explain an agent’s intentional action, then the reason involved must be an internal one, because this reason could neither produce nor explain the action if it was not related to this agent’s *S*. And, if it is necessary for the reason to be related to the agent’s *S*, then there is no sense in which the reason is external. For Williams, all reasons for action must have an explanatory dimension; but for this, all reasons must be related to an agent’s *S*, and any reason that must be related to an agent’s *S* is an internal reason and not external one.

Now, these four formulations of Williams’ argument are not the only ones available in the literature.<sup>55</sup> They simply happen to be the ones that I think best capture the basic spirit and line of thought behind Williams’ argument. Each of them, despite their minor faults, display the basic tie made between reasons and explanations, and explanations and motivations. In this way, external reasons are excluded, because external reasons lack a necessary connection to an agent’s *S*. At this point, we are ready to turn to the first objection to Williams’ argument.

### **§5. Shafer-Landau’s Objection**

Shafer-Landau rejects Williams’ conclusion on the grounds that his argument contains an equivocation on the term “reason.”<sup>56</sup> More specifically, Shafer-Landau

---

<sup>55</sup> For an extended discussion on the various formulations of Williams’ argument, see: Stephen Finlay, “The Obscurity of Internal Reasons,” *Philosophers’ Imprint* 9, no. 7 (July 2009).

<sup>56</sup> Russ Shafer-Landau, *Moral Realism*, 178-180. For other versions of the objection, see: David Brink, G. F. Schueler, E. J. Bond, and Derek Parfit. David Brink, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989), 62. G.F. Schueler, *Desire: Its Role in Practical Reason and the Explanation of Action* (Cambridge, Massachusetts: MIT Press, 1995), 60-61. E. J. Bond,



contends, premises 1 and 2 of Williams' argument, as I have presented them, concern "motivating reasons" while the conclusion concerns "justifying reasons." According to Shafer-Landau, if Williams had meant to conclude something about motivating reasons, the argument would have been sound albeit weak, since it could have no implications for justifying reasons.<sup>57</sup> But given that the conclusion is meant to be a thesis about justifying reasons (as Williams himself admits),<sup>58</sup> it follows that the argument is fallacious due to the shift in meaning.

Very roughly, a motivating reason is a reason that explains an action, and a justifying reason is a reason that justifies or legitimizes an action. Often this distinction is mirrored by the difference between reasons and motives. We typically say that a woman who killed her husband for money had a motive but not a reason to kill him; and, in this same instance, we say that the woman had a reason *not* to kill him (because it is wrong to do so), but this reason could not explain what she did, because she in fact acted contrary to it. Despite this distinction, it is possible for a justifying reason to be a motivating reason and vice versa, but a justifying reason need not explain an agent's action, because an action can be justified without an agent ever acting on it (hence, there would be no explanation of the action), and a motivating reason need not justify an agent's action, because an action can be explained without it being justified, e.g. the woman's desire for money explains but does not justify the murder of her husband.

---

*Reason and Value* (Cambridge: Cambridge University Press, 1983). Derek Parfit and John Broome, "Reasons and Motivation," *Proceedings of the Aristotelian Society* 71 (1997).

<sup>57</sup> Shafer-Landau, *Moral Realism*, 178-179.

<sup>58</sup> Williams, *Moral Luck*, 102-103.

Now Williams' argument runs into trouble, Shafer-Landau argues, because it fails to respect this distinction. More specifically, Shafer-Landau says:

The confusion lies in using the same term, 'reason', for two distinct kinds of thing: motives, which are psychological states, and genuine (good, justifying) reasons, which are not. Considerations to do with explanation of action are relevant to motives. [Reasons] Internalism is a thesis about justifying reasons. Thus [Williams'] argument will not do.<sup>59</sup>

According to Shafer-Landau, motivating reasons, as used in premises 1 and 2 from Williams' argument, are certainly tied to action explanation, because they are by definition dependent on our subjective motivational sets. But, Shafer-Landau insists, Williams' conclusion is meant to be a general statement about what reasons for action agents could have at all, including justifying reasons. This is problematic, Shafer-Landau suggests, because justifying reasons need not be connected to our subjective motivational sets. For, as we have seen, justifying reasons need not explain an agent's action.

I think Shafer-Landau's basic point here is that we should not and cannot limit what good reasons an agent can have simply on the basis of what motivating reasons this agent can have. Just because an agent is not motivated to act on a good reason, it does not mean the agent does not have this good reason. We normally think, for example, that even if a greedy person is not motivated to give money to charity, this person nevertheless has at least some reason to do so; for, we normally think that giving to charity would be a good or justified thing for him to do, regardless of what he wants to do. So, given the disparity of Williams' use of the term "reason," Shafer-Landau

---

<sup>59</sup> Shafer-Landau, *Moral Realism*, 178-179.



suggests, Williams' conclusion does not follow from the premises. Nothing in particular about the nature of justifying reasons can be drawn solely from considerations about the nature of motivating reasons. Shafer-Landau concludes, therefore, that Williams' argument is invalid.

## **§6. Reply to Shafer-Landau**

Shafer-Landau's objection is initially compelling, but there are a couple of things that can be said to vindicate the validity of Williams' argument. First of all, it is uncharitable to imply, as Shafer-Landau does, that Williams fails to respect the distinction between motivating reasons and justifying reasons. In his discussion of Williams' argument Shafer-Landau says, "Not every good reason is acted on and not every action is prompted by a good reason."<sup>60</sup> Clearly, however, Williams would not disagree, and it is evident from Williams' notion of a sound deliberative route that his position can accommodate precisely this point.

Recall Williams' example of the man who desires to drink the liquid in the bottle, because he believes it contains gin, when in fact it is petrol. Using this example, Williams aims to show generally that not every instance of an agent having a motivation to  $\phi$  is an instance of that agent having a reason to  $\phi$ . Since, according to Williams, there is no sound deliberative route from the man's  $S$  to his conclusion to drink the liquid, it follows that, despite the man's desire, there is no reason for him to do so. To be sure, the man is clueless that he has no reason, but it certainly seems sensible to say, as Williams does, that especially by the man's own lights it is true that he has no such

---

<sup>60</sup> Ibid., 178.

reason to drink the liquid. Next, recall my example of poor Franklin who has a reason to see the doctor, because, unbeknownst to him, there is a sound deliberative route from his *S* to his doing so. What this example shows is that, on Williams' view, it is possible for an agent to have a reason to  $\phi$ , but not a motivation to  $\phi$ , because he is unaware that there exists a sound deliberative route from his *S* to his  $\phi$ -ing, which grounds his reason to  $\phi$ . These two examples show that Williams thinks an agent can have a motivation without having a reason, and conversely a reason without a particular existing motivation.<sup>61</sup> The only thing I think we should conclude from this is that Williams not only respects, but consciously maintains, a distinction between motivating reasons and justifying reasons.

The most significant flaw in Shafer-Landau's objection, however, is that he fails to appreciate the central insight of Williams' argument: all reasons for action must have an explanatory dimension, i.e. the potential to explain an agent's action performed for that reason. According to Williams, whether we identify a given consideration *R* as a motivating reason or a justifying reason, it must be the case that, if *R* is *A*'s reason for acting on some particular occasion, then *R* must figure in a correct explanation of that agent's action. Shafer-Landau appears to gloss over this step in the sequence of Williams' argument when he accuses Williams of confusing the two types of reasons. To claim simply that there is a difference between motivating reasons and justifying

---

<sup>61</sup> This point needs to be treated with some care, since on Williams' view a reason may exist without an existing motivation, though not without at least some possible motivation, i.e. one that an agent could arrive at through a process of sound deliberation from her other existing motivations.



reasons is not enough to demonstrate a fault in Williams' argument, because Williams does not flatly conflate the two. To clarify this, Williams says:

It must be a mistake simply to separate explanatory and normative reasons. If it is true that *A* has a reason to  $\phi$ , then it must be possible that he should  $\phi$  for that reason; and if he does act for that reason, then that reason will be the explanation of his acting. So the claim that he has a reason to  $\phi$  – that is, the normative statement: “He has a reason  $\phi$ ” – introduces the possibility of that reason being an explanation...<sup>62</sup>

For Williams, premise 1 does not merely concern motivating reasons, but justifying reasons as well. If we accept that an agent could act for some justifying reason *R*, then, according to Williams, we must accept that *R* could be the explanation of that agent's action. Hence, all justifying reasons, insofar as they are considerations for which agents could act, must have an explanatory dimension. For this reason, justifying reasons (in addition to motivating reasons) are subject to the explanatory constraint of premise 2, and therefore must be related to our subjective motivational sets. Given that premise 1 is meant to cover any kind of reason for action, it follows that the term “reason” in Williams' argument does not shift meaning as Shafer-Landau suggests.

It seems to me that Shafer-Landau's objection arises only because he fails to appreciate the importance of premise 1 of Williams' argument. Once we see premise 1 in the correct light, it is clear that Williams does not confuse justifying reasons with motivating reasons, but in fact refines our understanding of reasons for action in general. With these considerations in mind, I think it is reasonable to conclude that

---

<sup>62</sup> Williams, *Making Sense of Humanity*, 38-39.

Williams does not commit a fallacy of equivocation, and, thus, that his argument is not invalid. To defend the validity of Williams' argument is one thing; it is quite another to defend its soundness. The next chapter will be devoted to precisely this task.



## CHAPTER 2: THE SOUNDNESS OF WILLIAMS' ARGUMENT

In the last chapter we introduced Williams' argument against external reasons. We also encountered one objection to the argument's validity, but concluded that it was unsuccessful. To my knowledge there are no other challenges to the argument's validity, so I will turn now to the argument's soundness. To be sure, the opponent of Williams' argument need only falsify one premise to demonstrate that it is unsound; given the relative plausibility of each premise, we might expect that this would be the most common approach. Indeed, this is borne out by the literature – none of Williams' critics, with regard to the premises as I have presented them, are willing to give up all three. Needless to say, however, we may exclude from consideration the truth of premise 3, since it merely expresses a definitional point about external reasons.

Before we begin, let's refresh our understanding with a quick look at Williams' argument once more. Williams claims that reasons, to be reasons, must be able to explain actions. But to explain actions, reasons must be related to subjective motivational sets. Since external reasons are by definition unrelated to subjective motivational sets, it follows that external reasons alone could never explain actions. Given this, Williams concludes, external reasons are no reasons at all; we are simply in error in believing that there are such things as external reasons. For convenience, I will restate the formalized version of the argument here:

- (1)  $R$  is a reason for  $A$  to  $\phi$  only if  $R$  could explain  $A$ 's  $\phi$ -ing.
- (2)  $R$  could explain  $A$ 's  $\phi$ -ing only if  $R$  is related to  $A$ 's subjective motivational set.
- (3) External reasons are not related to  $A$ 's subjective motivational set.
- (4)  $R$  is not an external reason for  $A$  to  $\phi$ , i.e. there are no external reasons for  $A$  to  $\phi$ .

### §1. Premise 1

As we have seen, premise 1 states a necessary condition for being a reason for action: some consideration  $R$  could not be a reason for  $A$  to  $\phi$  unless  $R$  could explain  $A$ 's  $\phi$ -ing. Recall that premise 1 does not require of a reason that it actually explain an agent's actions, because it is possible for an agent to have a reason to  $\phi$  even if she never in fact  $\phi$ -s for that reason. Again, perhaps Marsha has a reason to plant an apple tree, but decides she has better reason not to do so. In this case, Marsha's original reason was not in fact acted for, so the reason will not explain her action, but it could have, if she had acted for it, or will act for it in the future.

Premise 1 seems to me to be a relatively uncontroversial claim, given its acceptance among many of Williams' most devoted critics.<sup>63</sup> In fact, we might aid our understanding of premise 1 by looking at a couple of different ways other philosophers have expressed the same idea. Jonathan Dancy, an outspoken critic of Williams' argument, grants premise 1, and even affords it the status of a maxim: "A reason must be something for which someone could have acted, and in any case, where someone

---

<sup>63</sup> See for example: Jonathan Dancy, "Why There is Really no Such Thing as the Theory of Motivation," *Proceedings of the Aristotelian Society* 95 (1995); Korsgaard, "Skepticism about Practical Reason"; Heuer, "Reasons and Actions and Desires."



does act for that reason, the reason contributes to the explanation of her action.”<sup>64</sup> I think Dancy’s expression of premise 1 is helpful, because it makes clear an important presupposition at play in Williams’ argument. For Dancy, as for Williams, reasons must be able to explain an agent’s actions, because reasons must be the kind of thing for which an agent could possibly act. How could something be a reason for me to  $\phi$ , if I could not possibly  $\phi$  for that reason? And, if I could  $\phi$  for that reason, then that reason must be potentially explanatory of my action, because if I do *in fact* act for that reason, then that reason will be the explanation of my action. Hence, reasons for action must always have an explanatory dimension. This line of reasoning seems natural enough, but as we will see the basic presupposition on which it depends, i.e. that “a reason must be something for which someone could have acted,” is not universally accepted.

Stephen Finlay, a critic closer to Williams’ view, reads premise 1 in a slightly different manner. For Finlay, premise 1 is tantamount to a conceptual truth: “...the very conceptual content of “a reason for action” involves the concept of an explanation. To believe that *R* is a reason for action *is nothing other* than to believe it is a special kind of explanation of action.”<sup>65</sup> It is unclear to me whether treating premise 1 as a conceptual truth or not makes a significant difference to Williams’ argument. Presumably an objection that could falsify Dancy’s version could falsify Finlay’s just as well. In any case, I think Finlay’s expression helps to get at Williams’ basic point that it makes little sense to talk about a reason for action that could not, at least potentially, be the explanation

---

<sup>64</sup> Dancy, “Why There is Really no Such Thing as the Theory of Motivation,” 4.

<sup>65</sup> Finlay, “The Obscurity of Internal Reasons,” 13.

of an agent's action. Williams says, "If something can be a reason for action, then it could be someone's reason for acting on a particular occasion, and it would then figure in an explanation of that action."<sup>66</sup>

Despite the intuitive plausibility of premise 1, it is certainly not without detractors. In fact, there is a prominent style of objection that attempts to falsify premise 1 by means of counterexample.<sup>67</sup> To be clear, premise 1 would be falsified if it could be shown in at least one case that two things hold: (1) *R* is a reason for *A* to  $\phi$ , but (2) *R* could not explain *A*'s  $\phi$ -ing. Rachel Cohon, for one, offers the materials to build just such a counterexample.<sup>68</sup> According to Cohon, it is possible for an agent to have a reason to  $\phi$ , and for this reason to be incapable of explaining the agent's  $\phi$ -ing, because an agent may have various "psychological blocks" that prevent the agent from ever being able to  $\phi$ . Cohon says: "Such a block may occur where a phobia comes into play, such as a morbid fear of flying, or of germs or snakes, that always freezes the agent in her tracks on the verge of acting for the...reason."<sup>69</sup> Cohon herself does not offer a concrete example of such a case, but it is not difficult to construct one on the basis of what she says here.

---

<sup>66</sup> Williams, *Moral Luck*, 106.

<sup>67</sup> For objections related to this type see: Rachel Cohon, "Internalism about Reasons for Action" *Pacific Philosophical Quarterly* 74, vol. 4 (December 1993), 271; Millgram, "Williams' Argument Against External Reasons," 197-220; Shafer-Landau, *Moral Realism*, 180.

<sup>68</sup> It is unclear whether Cohon herself intends for the ideas noted here to constitute an objection to premise 1. Regardless of her intentions, I think that can be made to do so. Cohon, "Internalism about Reasons for Action", 270-271.

<sup>69</sup> *Ibid.*, 271.



Imagine that Tommy has recently located his long lost brother in another country, but that he must fly to this country to meet him. Let's stipulate here, however, that Tommy has a morbid fear of flying so severe that he is rendered physically incapable of getting on a plane. In this case, we can see that although Tommy has a reason to fly to this other country, this reason cannot explain Tommy's action, because he is incapable of acting in this way for it. Hence, there is a reason for Tommy to fly to another country, but this reason could not explain Tommy's doing so, because Tommy cannot in fact fly for the reason.

Does this type of example show that premise 1 is false? It may depend on how we are meant to interpret the severity of Tommy's situation. On the one hand, if we are meant to suppose that Tommy simply will never get on a plane because he finds it especially difficult to overcome his fear of flying, then it seems to me that such an example poses no serious problem for premise 1. Just because Tommy *will* never fly to another country for this reason, it does not mean that he *could* never do so for this reason. Presumably Tommy could, through various means, e.g. therapy, medication, etc., get over his fear of flying. And, if this is possible, then Tommy could in principle act for his reason. Assuming this is the case, it seems sensible to say that Tommy's reason still could in principle explain his action, even if he never actually overcomes his fear and gets on a plane.

On the other hand, if we are meant to suppose that it really is impossible for Tommy to fly to another country for this reason, then perhaps the example does pose a problem for premise 1. But how exactly are we to understand "impossible" here? How

would a psychological block, as Cohon describes, render someone completely incapable of doing something? Perhaps we could alter our example slightly to make the objection clearer. Suppose that, regardless of Tommy's fear, due to a worldwide catastrophe it comes to happen that there are no airplanes left on the face of the planet. In this case, Tommy really would be incapable of flying for this reason, because there would be no airplanes left for anyone to fly on, let alone Tommy. The pressing question then is this: could Tommy's reason ever explain Tommy's flying? Without much thought, it seems that we must answer in the negative – Tommy's reason could never explain his flying, because Tommy really would be incapable of flying to another country; hence, Tommy's action would and could never call for an explanation.<sup>70</sup>

Does this stronger example show that premise 1 is false? I think not, but ultimately it depends on the truth of what I have called the "basic presupposition" of premise 1. We have already seen Dancy's expression of this notion, Williams' version is not far off: "If it is true that A has a reason to  $\phi$ , then it must be possible that he should  $\phi$  for that reason..."<sup>71</sup> The basic idea here is that one cannot have a reason to perform an action unless one is able to perform that action at all. Unfortunately, Williams and others offer little, if any, direct support for this claim. This is problematic because, as

---

<sup>70</sup> One might question whether this new example really is all that different from the first, because just as Tommy *could* get over his fear of flying, he *could*, for instance, attempt to build an airplane from scratch, in which case it would still be in principle possible for Tommy to fly to another country to see his brother. Indeed, this seems to me to be on point. However, given that multiple critics (Cohon, Shafer-Landau, and Millgram) think that this type of counterexample can be made to work, for the sake of argument, I want to grant that Tommy's new case truly represents a challenge to premise 1. In this way, we will see that even in its strongest form, this type of objection fails. For references, see note 65 above.

<sup>71</sup> Williams, *Making Sense of Humanity*, 39.



we will see, the truth of premise 1 directly hinges on whether or not this claim is true, and not everyone thinks it is. So, unless Williams' position on this question can be defended, counterexamples like Tommy's case will threaten the truth of premise 1, and we will thus have to cast suspicion on the soundness of Williams' argument.

## §2. Reasons and Impossible Actions

Following Bart Streumer, I will define and talk about Williams' position on reasons and impossible actions in the following manner:

**(R):** "There cannot be a reason for a person to perform an action if it is impossible that this person will perform this action."<sup>72</sup> (In our terms:  $R$  is a reason for  $A$  to  $\phi$  only if  $A$  is capable of  $\phi$ -ing at all).

Before we assess the truth of (R) itself, let's get clearer about the logical relationship it bears to premise 1. First suppose (R) is false. In this case, it is possible for an agent  $A$  to have a reason  $R$  to  $\phi$  even if she cannot  $\phi$  at all. But, if it is impossible that  $A$  will  $\phi$ , then  $R$  could never explain  $A$ 's  $\phi$ -ing, because it will never be the case that  $A$ 's  $\phi$ -ing calls for an explanation. Assuming this is so, however, then it is possible that (1)  $A$  to has a reason to  $\phi$ , but (2)  $R$  could never explain  $A$ 's  $\phi$ -ing. But this is exactly the possibility we said would falsify premise 1. Hence, assuming (R) is false, it follows that premise 1 is false. This means that the truth of (R) is a necessary condition for the truth of premise 1, i.e. premise 1 cannot be true unless (R) is true. Now, this also means that, if (R) is true, the truth premise 1 cannot be eliminated outright. For, suppose (R) is true;

---

<sup>72</sup> For Streumer, "it is impossible that  $A$  will  $\phi$ " means "there is [no] nomologically and historically accessible or close possible world in which...[ $A$ ] will...[ $\phi$ ]." Streumer further clarifies that "A possible world is nomologically and historically assessable if and only if its laws of nature and its past are the same as those of the actual world." Throughout this chapter, I will follow this restricted sense of possibility. Bart Streumer, "Reasons and Impossibility" *Philosophical Studies* 136, 351-384 (2007), 358; Bart Streumer, "Reasons, impossibility and efficient steps: reply to Heuer" *Philosophical Studies* (June 2009), 2.

in this case, for any reason  $R$  that  $A$  has to  $\phi$ , it will always be possible for  $A$  to  $\phi$  for  $R$ . And, if it is always possible for  $A$  to  $\phi$  for  $R$ , then, it seems, it will always be the case that  $R$  could explain  $A$ 's  $\phi$ -ing, because, if  $A$  in fact  $\phi$ -s for  $R$ , then  $R$  will be the explanation of  $A$ 's  $\phi$ -ing.

From these considerations we can see that without the support of (R), premise 1 cannot be maintained. Thus, in order to save the truth of premise 1, we must save the truth of (R). And, although it is worthwhile to note that the truth of (R) does not absolutely guarantee the truth of premise 1, given the fact that challenges to (R) are the only way in which Williams' critics attempt to falsify premise 1, it seems fair to say that if a strong case can be made for (R), the presumption will fall in favor of premise 1. In other words, supposing we find no good reason to reject (R), and we find at least some good reason to accept (R), then we will have no good reason to think premise 1 is false.

But haven't we already seen that (R) is false? In our modified counterexample, we saw that although Tommy has a reason to fly to another country, he is incapable of flying for this reason, because there are no airplanes left to fly on at all. If something like this can be the case, then it looks like (R) must be false. However, I believe there are grounds for thinking that such counterexamples are unfair to begin with. In setting up our example we merely stipulated that Tommy had a reason to do something he is completely incapable of doing, but one might be suspicious of the intelligibility of such a stipulation. Does it really make sense to talk about someone having a reason to perform an action he or she simply cannot perform? Bare intuitions on this matter seem to be of no help, since different people will have different intuitions about cases



like Tommy's. For some, it will seem very strange to deny (R), because, unless there are other restrictions, there appears to be nothing to stop us from attributing an infinite number of, what Streumer calls, "crazy reasons" to ourselves and others.<sup>73</sup> To borrow an example from Streumer, it looks as though we would have to accept innumerable cases like the following:

Suppose that Jane is a person living in the twenty-first century. Given how bad the crusades, slavery and the two world wars were, there are reasons for Jane to prevent the crusades, slavery and the two world wars, even though, to do this, she would have to travel back in time and single-handedly change the course of history.<sup>74</sup>

Do we really think that Jane has a reason to travel back in time, even though she cannot do so?<sup>75</sup> Personally, I agree with Streumer; answering "yes" here simply sounds crazy.<sup>76</sup> For some, however, cases like these will not be intuitively problematic in the least. One might well respond, "just because something sounds crazy, it does not mean it isn't true – maybe Jane does have a reason to travel back in time." Initially, it is difficult to see how we might go about resolving the clash of intuitions here. Yet, as I will argue, between one of Streumer's more compelling arguments for (R), and some related ideas available in Williams' work, we have better reason to accept (R) than not.

---

<sup>73</sup> Streumer, "Reasons and Impossibility," 358.

<sup>74</sup> Ibid., 358-362.

<sup>75</sup> For this case, we can assume that Streumer has in mind nomological possibility. That is, for Streumer, time travel is impossible for Jane, because she belongs to a world in which the laws of nature prevent the possibility of traveling back in time.

<sup>76</sup> I should note that this is not the only point of Streumer's "argument from crazy reasons." According to Streumer, (R) is the simplest and least ad hoc explanation of the non-existence of crazy reasons. Streumer, "Reasons, impossibility and efficient steps: reply to Heuer," 5.

Moreover, given that there are no arguments for not-(R), but merely intuitions, I think we should side Williams, and reject that examples like Tommy's case falsify premise 1.

### §3. Two Arguments for (R)

The most compelling argument Streumer puts forward in favor (R) begins with a fairly plausible assumption about the purpose of practical deliberation. For Streumer, "when rational deliberation gives rise to the belief that we have most reason to perform a certain action, it should result in our trying to perform this action, (which may, if we succeed, coincide with actually performing this action)."<sup>77</sup> In other words, according to Streumer, when someone deliberates about what to do, we expect that the result of this deliberation will be the person trying to perform the action she has best reason to perform.<sup>78</sup> Why else would someone stop to consider what she had best reason to do? So, for example, when Olivia deliberates about how best to get to the store, and she determines that she has the best reason to ride her bike, we expect that, barring intervening factors, Olivia will try to ride her bike to the store. If Olivia's deliberation did not result in this, we might be inclined to ask "what was the point of her deliberation in the first place?" A difficulty arises, according to Streumer, when we see that not-(R), i.e. the claim that someone *can* have a reason to perform an action even if it she cannot perform that action, ultimately clashes with this assumption. Streumer's argument for this claim begins as follows:

---

<sup>77</sup> Ibid., 7.

<sup>78</sup> For Streumer, we should merely expect that someone *tries* to perform such an action, because it may be unreasonable in certain cases to expect that she will actually succeed in performing this action.



Suppose that (R) is false...In that case, whenever a person engages in deliberation about what to do, this person will not only have to judge which reasons there are to perform actions that he or she can perform, but will also have to judge which reasons there are to perform actions that he or she cannot perform.<sup>79</sup>

This is problematic, Streumer argues, since it will often be the case that the action we have best reason to perform is also the action that we simply cannot perform. So, imagine that Jane from before is deliberating about what to do in a given case and she happens to consider the horrors of past events. According to Streumer:

Given the enormous amount of suffering that was caused by the crusades, slavery and the two world wars, [Jane] will then almost always have to conclude that there is most reason for [her] to travel back in time and prevent the crusades, slavery and the two world wars.<sup>80</sup>

In line with our assumption about the purpose of practical deliberation, it follows from this that Jane must actually try to travel back in time to prevent these events, even though it is pointless for her to do so (because it is impossible for her to actually do so). Thus, according to Streumer, “if (R) is false, rational deliberation will almost always result in a person pointlessly trying to perform actions that he or she cannot perform,” but, of course, we do not think practical deliberation should have such results; hence, Streumer concludes, (R) must be true.<sup>81</sup>

---

<sup>79</sup> Streumer, “Reasons and Impossibility,” 365.

<sup>80</sup> Of course, this will not always be the case for every agent, but this would be irrelevant to Streumer’s point. For Streumer’s point to go through, this type of deliberative result need only happen for some agents some of the time. So, in this case, we may assume that especially by Jane’s own lights, she has best reason to travel back in time. Ibid., 365.

<sup>81</sup> Ibid., 365.

One might object to Streumer's argument by denying that his assumption about the purpose of practical deliberation is necessarily at odds with not-(R). In other words, one could grant that the purpose of practical deliberation is to issue in trying to perform certain actions, but not thereby give up the idea that people can have reasons to perform actions that they are incapable of performing.<sup>82</sup> Perhaps, for instance, one could simply note that when deliberating about what to do, and when trying to perform the actions one has best reason to perform, one should take into account *only* those reasons for which one is able to act. In other words, as a general rule, we could simply ignore or disregard the reasons we have to perform actions that we simply cannot perform. Thus, as the objection might go, Streumer's argument fails to rule out the possibility of the existence of reasons for impossible actions, and therefore fails to establish the truth of (R).

This kind of objection has some initial pull, but ultimately I think it does little to help the case for not-(R). If, in order to save Streumer's assumption about the purpose of practical deliberation, we are instructed to ignore or disregard the reasons we have to perform actions we cannot perform, it seems to me that this objection amounts to nothing more than the following recommendation: pretend as though such reasons do not exist at all. Yet, if we are meant to pretend that such reasons do not exist at all (because they make no difference to our deliberations or actions), it is unclear to me why we should think that there are any such reasons in the first place. At most we are left with the (not universally held) intuition that, despite the complete irrelevance such

---

<sup>82</sup> The basic idea for this objection comes from Streumer. Ibid., 366.



reasons have to our lives, we might nevertheless have them. But this is no argument, only a restatement of the original intuition for not-(R). So, unless something else can be added to this original intuition to help support not-(R), we have better reason to side with Streumer's argument and accept that (R) is true.

In addition to Streumer's argument, there are some relevant ideas available in Williams' work that can be utilized to help make the case for (R). Williams himself does not use these ideas to support (R), but despite this, I think they are applicable to the issue at hand. According to Williams, there is close relationship between failing to act on a reason and being blameworthy in some respect.<sup>83</sup> For example, when Charlie recognizes that he has a reason to help his grandmother move (e.g. because no one else is available to do so), and he has no better reason not to help her, and it is within Charlie's power to do so, and he believes that it is in his power to do so, but Charlie fails to act for this reason (e.g. because he carelessly forgets to wake up on time, or because he is too tired to help the day of the move, etc.), we typically think that Charlie is an appropriate target for blame, that is, we think Charlie is blameworthy in some respect.<sup>84</sup> One might ask Charlie, "you saw that you had good reason to help you grandmother, didn't you? Why didn't you help her?" Of course, by hypothesis, we know that Charlie

---

<sup>83</sup> Williams, *Making Sense of Humanity*, 41.

<sup>84</sup> To be clear, "blameworthiness" need not be taken in a moral sense. Williams provides the following example: "...one partner to a bank robbery may ruin it by an idiotic mistake and be blamed by his companion for the fact that they are in jail. His companion is not invoking the system of morality and does not think that this was an offence against moral canons. Nevertheless, there is something in the idea that the failure that is being blamed has some kind of ethical dimension to it." Williams, *Making Sense of Humanity*, 40.

has no good answer to this question, because he simply failed to act for a good reason he knew he had.

However, for Williams, the situation changes if we imagine that Charlie was in fact unable to help his grandmother move, say, because he is a quadriplegic. For the sake of argument, suppose that Charlie, in this modified case, still has a reason to help his grandmother move (again, because there is no one else to help her). Supposing this is so, when Charlie fails to help his grandmother for his reason (because he is physically incapable of helping his grandmother move), we would not think that he is blameworthy in any way. How could we hold him accountable for not helping her for this reason, when he was unable to do so at all?<sup>85</sup> Yet, if we cannot blame Charlie for not acting for a reason that he cannot act for, according to Williams, we seem to be denying that such a reason was ever a genuine option for Charlie in the first place.<sup>86</sup> But if this is so, then it seems to me that we have no basis for saying that Charlie ever had such a reason. By *reductio*, of course, this means that our original supposition, i.e. that Charlie had a reason to help his grandmother despite his condition, must be false.

So, the simple fact that we would not blame Charlie here (or Jane for not acting for her alleged reason to travel back in time to prevent past horrors) gives us good

---

<sup>85</sup> We may agree, I think, that there are some things Charlie could do to help his grandmother move even if he is physically immobile, but the action of interest in this case is physically helping her move, e.g. lifting boxes, furniture, etc., an action that Charlie is, by hypothesis, unable to perform.

<sup>86</sup> Williams makes this point by noting the apparent close connection between blaming after the fact and advising before the fact. If we blame A for not  $\phi$ -ing, then before A's inaction, it must have been appropriate for us to advise A to  $\phi$ , and conversely, if it would not have been appropriate for us to advise A to  $\phi$ , then it must not be appropriate to blame A for not  $\phi$ -ing, if A fails to  $\phi$ . In short, Williams says, "If 'ought to have' in the mode of blame corresponds to 'ought to' in the mode of advice, this strongly suggests that what it refers to is the agent's having (having had) a reason: 'ought to have' will carry the thought that the agent had a reason to act in the desired way but failed to do so." *Ibid.*, 41.



reason to deny that people can have reasons to perform actions that they simply cannot perform. Blame is never appropriate when people fail to act for reasons they cannot act for at all.<sup>87</sup> Thus, unless we are given some basis for thinking that people can have reasons to perform actions they cannot perform *despite* the fact that we would not blame them for failing to act for these reasons – again, besides the original intuition for not-(R) – we have better reason to deny that there are such reasons from the outset, and, therefore, conclude that (R) is true.<sup>88</sup>

Between Streumer's argument from deliberation and Williams' ideas on reasons and blame, we have at our disposal a strong case for the truth of (R). To reconnect to the bigger issue, given that critics must assume not-(R) in order to produce counterexamples to premise 1, like Tommy's case, we should not allow such counterexamples even to get off the ground. Recall that Tommy's case, where Tommy allegedly has a reason to fly to another country but he cannot do so because every plane on earth has been destroyed, was intended to show that (1) someone could have a reason, but that (2) this reason could not explain this person's actions. Given that examples like this one cannot be produced without assuming not-(R), and that we now

---

<sup>87</sup> This is not to say that reasons do not exist *whenever* blame is inappropriate, because sometimes we cannot rightfully blame someone who failed to act for a reason she certainly had; sometimes intervening factors will excuse an agent from being blamed in this way. The claim here is simply that blame is not appropriate when someone fails to act for a reason that she was incapable of acting for in the first place.

<sup>88</sup> One might draw a parallel between the basic idea of this argument and the "ought-implies-can" principle. Certainly there would be something to this – we might even condense the point here to "reason-implies-can." It is unclear, however, how far this parallel might be able to go. For a discussion of the relationship between obligations and reasons, but as it specifically relates to Williams' work, see: Jonny Anomaly, "Internal Reasons and the Ought-Implies-Can Principle," *Philosophical Forum* 39, no. 4 (December 1, 2008).

have good reason to affirm (R), we must conclude that such counterexamples do not falsify premise 1. Moreover, since there are no other types of challenge to premise 1, the presumption rightfully falls in its favor. This means that if there is a problem with Williams' argument, it must be found elsewhere than with premise 1. On this note, we may turn to premise 2.

#### **§4. Premise 2**

As we saw in the first chapter, premise 2 states a necessary condition for being the explanation of an agent's intentional action: for any reason  $R$  to explain an agent's action,  $R$  must be related to the elements of an agent's  $S$ . Recall that from our discussion of premise 1, when an agent in fact  $\phi$ -s for some reason  $R$ ,  $R$  becomes the explanation of this agent's  $\phi$ -ing. Now, according to premise 2, such a consideration could not even become the explanation of an agent's  $\phi$ -ing unless it was already connected to at least one desire on the part of an agent. How could we explain someone's intentional action without citing something that could motivate, and so literally move, that person to act? According to the belief-desire theory of action explanation, the theory which Williams relies on to support premise 2, we simply cannot do this; reference to a motivational state, like a desire, is always a necessary piece of intentional action explanation.

To be clear, premise 2 does not require of a reason that it be related to an agent's *existing* motivational set, because it is possible for a reason to be potentially explanatory of someone's  $\phi$ -ing even if that person is not, or never will be, actually motivated to  $\phi$ . This brings us back to Williams' notion of a "sound deliberative route,"



something we encountered in discussing Williams' *reasons internalism*. Recall that a sound deliberative route is a path of reasoning that begins from an agent's existing motivations and progresses soundly to her practical conclusion to  $\phi$ . For Williams, given that such "processes of deliberation can have all sorts of effect on  $S$ ," one of which is the production of new desires, it follows that a reason can be potentially explanatory of an agent's  $\phi$ -ing so long as it is related to at least one desire that could be arrived at, from her existing desires, through the process of sound deliberation.<sup>89</sup>

To see what this might look like, let's play off of an old example. Suppose Franklin has a reason to see the doctor, because he has a life-threatening, undiagnosed disease. According to Williams, despite the fact that Franklin has no *existing* desire to see the doctor, because he is unaware of his condition, Franklin's reason to see the doctor still could explain his doing so, because, we can suppose, there is a sound deliberative route that could produce in Franklin a new desire to see the doctor. Of course, Franklin may never become aware of this sound deliberative route, and so never acquire the right desire to see the doctor, but, according to Williams, it does not follow from this that Franklin's reason could not, in principle, explain his going to see the doctor. To summarize this result, we should say: to be potentially explanatory of an agent's action, a reason must be either (1) related to an existing desire of the agent, or (2) related to one that could be arrived at through a sound deliberative route from her existing desires. Hereafter, when premise 2 says " $R$  must be related to an agent's  $S$ ," we will understand that this is shorthand for (1) and (2).

---

<sup>89</sup> Williams, *Moral Luck*, 105.

The basic idea behind premise 2 is not all that controversial. Many philosophers are willing to grant that reasons must somehow be tied to motivations to explain intentional actions. Christine Korsgaard, for example, says, "...unless reasons are motives, they cannot prompt or explain actions. And, unless reasons are motives, we cannot be said to be practically rational...Thus, it seems to be a requirement on practical reasons, that they be capable of motivating us."<sup>90</sup> The real controversy surrounding premise 2, as we will see, is not whether reasons must be tied to motivations to explain intentional actions, but over the nature of motivation itself. Of course, since Williams subscribes to the Humean theory of motivation (the HTM), he thinks the presence of a desire is a necessary condition for motivation, and so for explanation. But there are certainly philosophers who outright reject the HTM. Before we confront this wider issue, however, I want to consider a more direct objection to premise 2. This objection comes from Rachel Cohon, and it consists of numerous counterexamples intended to demonstrate that explanations of intentional actions need not reference an agent's desires. In showing why none of these counterexamples hit their mark, I hope to clarify, and perhaps partially vindicate, the claim of premise 2. Immediately following this, I will return to the wider issue of motivation – are desires the only thing that can motivate, and so explain, intentional actions? As I will argue, the truth of premise 2 depends on the truth of the HTM.

---

<sup>90</sup> Korsgaard, "Skepticism about Practical Reason," 11.



## §5. Cohon's Objection

Let's begin with the way Williams himself expresses premise 2. Williams says "nothing can explain an agent's (intentional) actions except something that motivates him so to act."<sup>91</sup> And, elsewhere he says, "When [a] reason is an explanation of [someone's] action, then of course it will be, in some form, in [this person's] S, because certainly – and nobody denies this – what [this person] actually does has to be explained by [this person's] S."<sup>92</sup> Recall that, for Williams, an agent's S is the complete collection of an agent's desires, where "desire" is a term of art meant to cover such things as commitments, loyalties, personal projects, etc.<sup>93</sup> Accordingly, for Williams, anytime we want to explain an agent's intentional action, our story must incorporate at least one desire on the part of that agent; otherwise, our explanation will be incomplete.

Cohon attacks this idea, but she begins with her own reading of premise 2. For Cohon, premise 2 states: "Intentional action is impossible, or at least inexplicable, in the absence of an element of S that the action serves or furthers."<sup>94</sup> On this reading, however, Cohon thinks premise 2 is false, because "many actions are done intentionally and may be readily explained, even though there is no element of the agent's S that the action serves or furthers."<sup>95</sup> Cohon attempts to demonstrate this claim with the help of

---

<sup>91</sup> Williams, *Moral Luck*, 107.

<sup>92</sup> Williams, *Making Sense of Humanity*, 39.

<sup>93</sup> Williams, *Moral Luck*, 105.

<sup>94</sup> Cohon, "Are External Reasons Impossible?" 549.

<sup>95</sup> Williams, *Moral Luck*, 549.

numerous counterexamples. Before we examine these alleged counterexamples, we should correct Cohon's reading of premise 2 so that it aligns better with Williams' intentions. This correction will not alter the intended effect of the counterexamples on premise 2.

Certainly Williams maintains that desires are necessary to explain intentional action, but Williams would deny that such desires must actually be *served* or *furthered* to do this work, as Cohon suggests. Call to mind Williams' example of the man who desires a gin and tonic, and believes the bottle in front of him contains gin, when in fact it is petrol.<sup>96</sup> About this case, Williams says, "...if [the man] does drink [the liquid], 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."<sup>97</sup> From this we can see that, for Williams, despite the fact that the man's original desire for a gin and tonic will not actually be served or furthered by drinking the contents of the bottle (because the bottle contains petrol and will likely injure the man when he drinks it), we can nevertheless explain the man's action with the help of his desire. We might explain: "The man drank the petrol because he desired a gin and tonic, but unfortunately he was unaware that the bottle did not contain gin." Of course, according to Williams, we cannot *justify* the man's action with this statement (because the man's action, even by his own lights, was a bad thing to do), but it does not follow from this that we cannot

---

<sup>96</sup> Ibid., 102.

<sup>97</sup> Ibid.



explain the man's action with the same statement.<sup>98</sup> Thus, according to Williams, while a desire is required to explain someone's intentional action, the desire need not be actually served or furthered to do so.

In view of this, we can see that Cohon's reading of premise 2 is too strong. With little effort though I think we can strike a compromise between Cohon's interpretation and Williams' intentions. Let's say now, echoing Cohon's original reading of premise 2, that: *A's intentional actions are impossible, or at least inexplicable, in the absence of an element of her S that the action serves or furthers, or A believes to serve or further.* This expression of premise 2 is slightly weaker than Cohon's original reading, but it is closer to Williams' view, because it allows for cases like the gin/petrol example. The man *believed* that by drinking the contents of the bottle he would serve or further his desire for a gin and tonic. Certainly the man was wrong about this, but this does thereby not hinder our ability to use the man's desire in an explanation of his action.

With this in mind, we may note that what Cohon, or anyone else, needs to show to falsify premise 2 is that desires in general are unnecessary to explain intentional actions, regardless of whether such desires are actually served or furthered by these actions. In the form of a counterexample, one would have to show that in at a least one case two things hold: (1) we can explain someone's intentional action, but (2) without incorporating a desire on this person's part to do so. Let's turn now to Cohon's proposed counterexamples to see if they can meet these two requirements.

---

<sup>98</sup> Williams, *Moral Luck*, 103.

Cohon produces a handful of ordinary examples, in part, I suspect, to show that we need not stretch ourselves very far to see that premise 2 is false. I will paraphrase just three of her examples:<sup>99</sup>

- (1) One evening a man packs a lunch for the next day like he always does, even though he knows he is supposed to meet a friend for lunch the next day.
- (2) Someone is given a posthypnotic suggestion, rises at a certain signal, gets a pencil and paper, and performs a calculation.
- (3) Someone is swayed by a candidate's personality and votes for him without thinking about whether he wants him in office (and it is stipulated that he would not).

About these examples, Cohon says first, "These actions are clearly done intentionally. The agent knows what he is doing and in some of the cases has to monitor his own behavior through a series of steps."<sup>100</sup> Secondly, Cohon says, all of these actions are explainable: (1) the man packed his lunch out of habit; (2) the woman performed the calculation due to the posthypnotic suggestion; and, (3) the man voted for the candidate because of the candidate's powerful personality.<sup>101</sup> Lastly, Cohon insists, these "...explanations do not mention any of the agents' desires, aims, projects, loyalties, or other things Williams allows in an S."<sup>102</sup> Thus, Cohon concludes, "it does not

---

<sup>99</sup> Cohon, "Are External Reasons Impossible?" 549.

<sup>100</sup> With regard to this point, Cohon is right to note: "These are not, of course, wise things to do. But some of our actions are not very intelligent. The fact that an action is not intelligent, however, does not imply that it is not intentional." Ibid.

<sup>101</sup> Ibid.

<sup>102</sup> Ibid.



seem to be true that nothing can explain an intentional action except an element of the agent's *S*..."<sup>103</sup>

At first blush, Cohon's counterexamples appear to falsify premise 2 rather easily. For, it seems that we have three clear-cut cases of (1) intentional action explanations that (2) do not incorporate any element of an agent's *S*. For the sake of argument, let's simply grant to Cohon that in each case, no desire is needed to explain the person's action. In other words, we may grant that Cohon's counterexamples meet the second requirement for the falsification of premise 2. We should not, however, agree that these counterexamples meet the first. This is because, it seems to me, Cohon equivocates on the term "intentional." In one sense, Cohon is absolutely right – all of these actions are intentional, in the sense that these people are consciously monitoring their behavior throughout the course of their actions. The woman in case (2), for instance, obviously has not *accidentally* performed the calculation in the way that someone might accidentally knock over a glass. But this non-accidental sense of "intentional" is not what Williams has in mind for premise 2. For Williams, to act intentionally is *to act for a reason*. Accordingly, the claim of premise 2 is *not* that we cannot explain intentional actions in the weaker, non-accidental sense, without including a desire, but rather that we cannot explain intentional actions in the stronger, acting-for-a-reason sense, without including a desire. We may well be able to explain intentional actions in Cohon's weaker sense without including a desire, but it does not follow from this that we can explain intentional actions in Williams' stronger sense

---

<sup>103</sup> Ibid., 549-550.

without including a desire. Due to this equivocation, Cohon's counterexamples fail to meet the first requirement for the falsification of premise 2, since they do not represent genuine cases of intentional actions, and therefore they fail to falsify premise 2.

One might question, however, whether Cohon must commit herself to the weaker sense of intentional action to make her counterexamples work. Perhaps it is true in case (2), for instance, that we must say the woman is acting intentionally in the weaker sense, i.e. not acting for a reason (because her action was the result of a posthypnotic suggestion). But perhaps in cases (1) and (3) the men could be made out to be acting intentionally in Williams' stronger sense: (1) the man packed his lunch *for the reason that* he believed he needed a lunch the next day (even though he forgot about his plan to meet his friend); and, (3) the man voted for the candidate *for the reason that* he found the candidate's personality especially powerful (even though he did not consider the long term consequences of his decision). If (1) and (3) can be altered in this way to meet the first requirement for the falsification of premise 2, then perhaps such counterexamples falsify premise 2 after all.

Ultimately I do not think such a response can save Cohon's counterexamples. This is because the new explanations of cases (1) and (3), while incorporating the stronger sense of intentional action, now seem to require the addition of a desire on the men's part. In case (1), it is not enough to say that the man packed his lunch because he believed he needed a lunch for the next day, because it is possible for someone else to have the exact same belief, but not pack a lunch. In order to explain why this man packed his lunch, we must include the fact that he also *desires* to have a lunch for the



next day. If the man did not have a desire like this, it is unclear to me why he, or anyone else, would pack a lunch at all. The same can be said for case (3); it is not enough to say that the man voted for the candidate because he found the candidate's personality especially powerful, since it is possible for someone else to have the exact same impression, but not vote for the candidate. In order to explain why this man voted for the candidate, we must include the fact that he, at least in that moment, *desired* the candidate to win. Without a desire like this, another person who also found the candidate's personality especially powerful might yet remain perfectly indifferent to whether this candidate gets elected, and as a result not vote for him.

So, the bottom line for Cohon this: if we grant that cases (1) and (3) meet the second requirement for the falsification of premise 2, because they now represent cases of Williams' stronger sense of intentional action, it seems to me that we must in so doing retract our initial concession that these cases meet the first requirement for the falsification of premise 2, because they now require reference to desires on the part of the men involved. Thus, given that Cohon's counterexamples meet either the first or second requirement for the falsification of premise 2, but not both, we must conclude that Cohon's counterexamples fail to falsify premise 2.

## **§6. Desires and Motivation**

Premise 2 may well be safe from this objection, but there are certainly other ways one might attempt to block the claim. Both Cohon and Ulrike Heuer, for instance, argue that the strength of premise 2 entirely depends on an unfair assumption Williams makes about the nature of motivation. When Williams states, "nothing can explain an

agent's (intentional) actions except something that motivates him so to act," Cohon thinks that Williams has unfairly helped himself to a particular interpretation of "motivation."<sup>104</sup> More specifically, Cohon says:

When we read "something that motivates" we naturally think of just this meaning: something that causes the person to act. Of course we cannot explain [someone's]... action without mentioning whatever it was that caused or impelled [this person]...so to act...But this fact carries few implications about the nature of whatever it was that caused him to act.<sup>105</sup>

In other words, according to Cohon, even if we admit that we cannot explain someone's intentional action without including something that motivated this person to act, we are not necessarily committed to Williams' claim that a desire must be the motivational piece of our explanation, because it might be that something other than a desire can motivate someone to act. Heuer, another critic of Williams' argument, makes a similar point:

And why should it be the case that only desires can explain actions?...It seems that the Humean trades on an equivocation: she first says that only motivationally relevant attitudes can explain actions, and then adds that 'desire' in its philosophical usage is just another word for 'motivationally relevant attitude.' But if so, it should not exclude any plausible candidate from being a 'motivationally relevant attitude.'<sup>106</sup>

Heuer's way of putting the alleged problem with premise 2 is helpful because it makes explicit the exact assumption Williams depends on to make the claim of premise 2 go through, namely, the Humean theory of motivation. In chapter 1, we briefly introduced the HTM as the view that beliefs are not sufficient, and desires are necessary, for

---

<sup>104</sup> Williams, *Making Sense of Humanity*, 107.

<sup>105</sup> Cohon, "Are External Reasons Impossible?" 551.

<sup>106</sup> Heuer, "Reasons and Actions and Desires," 54.



motivation. Heuer is right to note that Williams assumes the HTM, but I think she is wrong to imply that it is a groundless assumption. In chapter 1, I suggested that Williams falls in the camp of philosophers who maintain that desires are necessary for motivation, because, whereas beliefs are merely cognitive states that aim to reflect the world, desires are conative states that drive and impel agents to alter the world.<sup>107</sup> Despite the intuitive plausibility of the HTM, it is most certainly not a universally accepted view. But before we say anything more about the truth or plausibility of the HTM, it will be useful to get clearer about the logical relationship it bears to premise 2.

First, suppose the HTM is true. In this case, desires are necessary for motivation. Given that intentional action explanations require reference to an agent's motivations (a premise which even Williams' critics are willing to grant), it follows that intentional action explanations require reference to an agent's desires. So, if the HTM is true, then desires are necessary for the explanation of intentional actions. In essence, this is the claim of premise 2; thus, if the HTM is true, we have good reason to conclude that premise 2 is true as well. Now, in contrast, suppose the HTM is false, i.e. that desires are not necessary for motivation. In that case, it is entirely reasonable to assume that other psychological states, besides desires, can be responsible for motivating agents to act. But, if this is so, then desires are unnecessary to explain intentional actions. This contradicts the claim of premise 2; thus, if the HTM is false, then there is good reason to reject premise 2. From these considerations, we can see that the truth of premise 2, as both Cohon and Heuer suggest, depends on the truth of the HTM, and consequently, if

---

<sup>107</sup> See chapter 1, section 4.

there is a problem with the HTM, premise 2 can be plausibly blocked, and so the soundness of Williams' argument will be placed in serious jeopardy. Thus, to save premise 2 and Williams' argument, the HTM needs to be defended.

At this point we can turn to the HTM itself. I want to begin with what I take to be one of the most plausible alternatives to the HTM, a view I will refer to generally as *anti-Humeanism*. In essence, any version of anti-Humeanism will deny the central claim of the HTM, i.e. that desires are necessary for motivation. However, since there are several, complex ways one could spell out such a denial, I want to work with the most straightforward and attractive version I think is available. Despite that I will refer to this particular view as anti-Humeanism, it is worthwhile to note that there are many views, rather unlike this one, that could be appropriately called by the same name.<sup>108</sup> Following a brief description anti-Humeanism, I will present what I think is the most powerful argument against it. This argument comes from Michael Smith, a strong defender of the HTM, and we will see that, taken by itself, this argument also provides independent support for the HTM, and therefore for premise 2. To be sure, my aim is not to offer a decisive case for the HTM, as this would stretch us well beyond our present purposes. Rather, my aim here is simply to build a presumptive case for the HTM by dislodging one prominent alternative to the HTM itself. In the final section of this chapter, I will summarize the major results of Williams' argument so far.

---

<sup>108</sup> For a few examples of such views, see: Jonathan Dancy, *Moral Reasons* (Oxford: Basil Blackwell, 1993); John McDowell, "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian Society*, vol. 52, 13-29; Thomas Nagel, *The Possibility of Altruism* (Oxford: Oxford University Press, 1970); Thomas Scanlon, *What We Owe to Each Other* (Cambridge, Massachusetts: Harvard University Press, 1998).



## §7. Anti-Humeanism

As we have seen, the HTM says that no belief by itself is sufficient for motivation, because at least one desire is always required in addition to a belief to motivate an agent to act. It seems to me that the most compelling alternative to the HTM, i.e. anti-Humeanism, is the view that at least some beliefs by themselves are sufficient for motivation, namely, *moral* beliefs.<sup>109</sup> Even without the help of a desire, the anti-Humean contends, a moral belief is enough, by itself, to motivate an agent to act. This is because the content of a moral belief is special in a way that the content of a non-moral belief is not. Granted, a precise distinction between moral and non-moral beliefs is not perfectly easy to draw, but for our purposes I think we can rely on a few ordinary examples to help provide us with an intuitive grasp of the essential difference between the two.

Very roughly, then, a moral belief is a belief about the moral features of some act or state of affairs. For example, the belief that “stealing is wrong” is a moral belief, because it assigns a negative moral quality to the act of stealing. A non-moral belief, in contrast, is a belief about some state of affairs that simply lacks any moral features. We might say, for example, the belief that “there is a chair in Marsha’s room,” is a non-moral belief, because the state of affairs it purports to describe involves no moral features whatsoever; it simply describes the physical arrangement of a chair in a room.

---

<sup>109</sup> For the moment, we should assume *cognitivism* about moral beliefs to help get this view off the ground. Cognitivism is the view that moral claims express genuine propositions, that is, statements that can be true or false.

To return to the issue of motivation, the anti-Humean may concede that non-moral beliefs by themselves are insufficient for motivation, but she will not make the same concession about moral ones. Unlike the belief that “there is a chair in Marsha’s room,” for instance, which requires a relevant desire to help motivate an agent to act, e.g. the desire to sit down, the anti-Humean will insist, the belief that “stealing is wrong” can, all by itself, motivate an agent to act in some manner, e.g. to refrain from stealing. According to the anti-Humean, so long as the agent has an adequate understanding of the given moral proposition, and sincerely believes the proposition itself, this agent will be motivated to act in the appropriate manner. Of course, the anti-Humean need not say that a moral belief will always produce an overriding motivation to act in some way, but only that such a belief will produce at least some minimal motivation to do so. Perhaps, for instance, Marsha is typically motivated to refrain from stealing (because she believes that stealing is wrong), but recently other motivations have taken over, because she has fallen on hard times and cannot feed her family. In this case, the anti-Humean might say, it is not that Marsha is not motivated to refrain from stealing at all, but rather that she is so motivated, though not enough, in this case, to prevent her from actually stealing.

Characterized in this way, anti-Humeanism seems to me to be a fairly plausible view. In fact, I think we can agree with Shafer-Landau, an outspoken critic of the HTM, that this version of anti-Humeanism seems to line up best, or at least really well, with the testimony of ordinary people on the issue of their own motivations. According to Shafer-Landau, people often say things like, “I didn’t *want* to x (stand in front of that



bullet, accept the blame, remain chaste), but duty called. I knew what had to be done, and did it, desires (passions, wants, inclinations) be damned.”<sup>110</sup> In other words, according to Shafer-Landau, despite that we sometimes have no desire to  $\phi$ , or, that we even have a strong desire not to  $\phi$ , we often report that we are nevertheless motivated to  $\phi$ , simply because we believe that  $\phi$ -ing is the right thing to do. On this point, I think Shafer-Landau is right; anti-Humeanism seems to capture well the way people ordinarily talk about at least some of their own motivations.

The ardent defender of the HTM, however, will likely be unimpressed by this consideration. She might respond, “Just because people ordinarily think they are motivated without the help of desires, it does not mean that they are – people are notoriously wrong about the inner workings of their own psychology.” Certainly there is something to this point, though I think the presumption might yet remain in the favor of anti-Humeanism. We should not, of course, think that just because a view is popular, it cannot be overturned by argument, because, at least in the case of anti-Humeanism, I think it can. As previously mentioned, I think Michael Smith’s argument not only provides strong reason to drop anti-Humeanism, but to pick up the HTM as well. At this point, we may turn to Smith’s argument itself.

## **§8. Smith’s Argument**

One might wonder why we need a special argument against the view that moral beliefs are sufficient to motivate. After all, people with the same moral belief(s) are often motivated differently. We can imagine, for example, that both Marsha and Olivia

---

<sup>110</sup> Shafer-Landau, *Moral Realism*, 123.

have the same moral belief, that keeping promises is the right thing to do, but whereas Marsha is motivated to keep her promises, Olivia is not. In explanation of this, Olivia might say, "Sure, I believe that promise keeping is the right thing to do, but I really don't much care to do the right thing." If this is possible, i.e. that two people share the same moral belief but where one is motivated the other is not, then it is plainly false to say that moral beliefs are sufficient for motivation.

Reasonable as this line of thought seems at first, as it stands, it is open to numerous attacks. The anti-Humean is likely to question first whether Olivia really is devoid of any motivation to keep her promises. Could it not be that Olivia really has some motivation, though quite minimal, which always happens to be overridden by other more powerful motivations? Or, if Olivia truly has no such minimal motivation, then perhaps we can question whether Olivia really understands what it means for an action to be *right*. And, even if these considerations fall through, it might be sensible to look for other relevant differences between people in actual cases of divergent motivation. Perhaps the normal motivation provided by a moral belief is stunted when someone has an abnormal psychological condition, e.g. being overly anxious, extremely fearful, depressed, etc. Thus, the anti-Humean might respond, just because two people with the same moral belief can be motivated differently, it does not mean that moral beliefs by themselves cannot motivate at all.

The anti-Humean seems to have a point here, but I think the original objection concerning Marsha and Olivia can be clarified and sharpened to make it much stronger. Recall that, from our initial discussion of the HTM, according to Smith and others, beliefs



and desires are distinct psychological states, because they have different “directions of fit.” Beliefs, Smith says, have a “mind-to-world” fit, because beliefs aim to fit the world, and when a given belief does not, it is the belief that must change.<sup>111</sup> If, for example, Olivia believes that there is coffee in her kitchen, and she discovers that in fact there is no coffee there, then her belief must change. Desires, on the other hand, have the opposite direction of fit, i.e. a “world-to-mind” fit, because, according to Smith, desires aim to make the world fit them, and when the world does not fit a given desire, it is the world that must change.<sup>112</sup> If Olivia desires that there is coffee in her kitchen, and she finds no coffee there, then it is the world that must change, not Olivia’s desire.

Now, our anti-Humean runs into trouble, Smith argues, when she is confronted with the question of which direction of fit the supposed motivationally powerful moral beliefs have. Do these moral beliefs have a mind-to-world fit like other beliefs, or a world-to-mind fit like desires? At least initially, this pressures the anti-Humean into an uncomfortable dilemma: either moral beliefs are belief-like or desire-like in their direction of fit. If moral beliefs are belief-like, then it seems the anti-Humean must say that moral beliefs aim to represent actual states of affairs, which is in itself a contested issue – what possible properties in the world could these beliefs be picking out?

As some would argue, it seems much more plausible to say that moral beliefs are in fact something more like projections of our own moral sentiments, and that they do

---

<sup>111</sup> Smith, *The Moral Problem*, 111-116.

<sup>112</sup> *Ibid.*, 115.

not aim to represent objective properties in the world at all.<sup>113</sup> Of course, this issue is far too complex for us to sort out here, but at the very least, given how thorny the issue is, I think it shows why someone might want instead to treat moral beliefs as having the other direction of fit. For some, this will seem to be the right approach anyway, given the apparent action-guiding character of moral beliefs. Moral beliefs, unlike non-moral beliefs, seem to draw or impel people to alter the way the world is. In this sense, moral beliefs are more desire-like in their direction of fit; if we recognize, for instance, that no one is helping an injured person in need, our moral beliefs seem to drive us to bring about some change in the world. Yet, if moral beliefs are in fact desire-like, and thus aim to make the world fit them, then perhaps it was a mistake to categorize them as beliefs in the first place. Indeed, this is precisely the suggestion offered by one prominent view: moral beliefs are not really beliefs after all, but instead express, or “emote,” non-cognitive states including, among other things, desires.<sup>114</sup> But clearly, if something like this is case, then anything unique about anti-Humeanism seems to vanish; a desire, or some desire-like state, would still be a necessary condition for motivation. Thus, with either direction of fit, the anti-Humean seems to run into trouble.

But there is an obvious alternative to the horns of this dilemma. Why could the anti-Humean not simply say that moral beliefs have *both* directions of fit? This not only

---

<sup>113</sup> For one prominent example of such a “projectivist” view, see: Simon Blackburn, *Spreading the Word* (Oxford: Oxford University Press, 1991).

<sup>114</sup> One striking example of such a view is Ayer’s “emotivism;” see: A. J. Ayer, *Language, Truth, and Logic* (New York: Dover Publications, 1952).



seems reasonable, but also seems to capture what the anti-Humean wanted to say in the first place. Moral beliefs seem to be a unique kind of state that both aim to fit the world, and, in some sense, aim to make the world fit with them. If a response like this is could be spelled out adequately, then perhaps Smith's objection poses no special problem for the anti-Humean. According to Smith, however, this kind of response is ultimately untenable.<sup>115</sup> Consider first what the anti-Humean would have to say about a single moral belief with the content *P*, held by some agent *A*. For Smith, if *A*'s belief that *P* has both directions of fit, then by definition, if *A* discovers that not-*P*, it will be both the case that *A*'s belief must change to fit the world, *and* the world must change to fit *A*'s belief.<sup>116</sup> Taken literally, Smith concludes, "...the idea that there may be a state having both directions of fit is just plain incoherent."<sup>117</sup>

Smith acknowledges, however, that the anti-Humean need not be committed to such an unsophisticated view. According to Smith, the more "subtle" anti-Humean might insist that a moral belief, e.g. a belief of the form " $\phi$ -ing is right," can have both directions of fit, and this is not incoherent, because such a belief, unlike other ordinary beliefs, has more than one content.<sup>118</sup> In other words, the anti-Humean could say, a moral belief always has two distinct contents, one of which belongs to its belief-like element, and the other to its desire-like element. The moral belief that "promise-

---

<sup>115</sup> Smith, *The Moral Problem*, 117-120.

<sup>116</sup> *Ibid.*, 118.

<sup>117</sup> *Ibid.*

<sup>118</sup> *Ibid.*, 118-119.

keeping is right,” for example, has a belief-like content, i.e. that in fact promise-keeping is right, but it also has a distinct desire-like content, something roughly like: promise-keeping must be brought about. In this way, the anti-Humean could argue, the different directions of fit in the same moral belief would never conflict, because they belong to the moral belief’s two distinct contents.

This suggestion seems both natural and appealing. Certainly we would not think, for instance, if Marsha had the belief that “ $\phi$ -ing is right,” and she discovered that  $\phi$ -ing is in fact *not* right, that the world must now change to fit Marsha’s belief, i.e. that Marsha must make it so that  $\phi$ -ing becomes right. In this sense, moral beliefs have the same direction of fit as other beliefs; they aim to fit the world, and must change when they do not. But, as the new suggestion goes, this does not mean that such beliefs cannot also have the other direction of fit. So long as Marsha’s belief has another, distinct content, e.g. that  $\phi$ -ing must be done, it can without difficulty have a desire-like direction of fit as well. So, the anti-Humean might ask, “Where’s the problem?”

For Smith, the problem with this suggestion is that, supposing a mental state does in fact have two distinct contents, we no longer have any reason to think that this mental state truly is a “*single, unitary, kind of state*.”<sup>119</sup> Smith’s argument for this is simple: “...it is always at least *possible* for agents who are in a belief-like state to the effect that their  $\phi$ -ing is right to none the less lack any desire-like state to the effect that they [must]  $\phi$ ...the two can always be pulled apart, at least modally.”<sup>120</sup> In other words,

---

<sup>119</sup> Ibid., 119.



according to Smith, if a mental state has two contents, then this mental state must really be two mental states, because its two contents can always come apart.<sup>121</sup> This poses an obvious problem for the anti-Humean: if it is always possible for the contents of a moral belief to come apart, then it is always possible for one person to be in its belief-like state but not its desire-like state, and vice versa. Yet, if this is possible, then the central claim of anti-Humeanism, i.e. that moral beliefs, by themselves, are sufficient for motivation, must be false. To return to our original example, while Marsha may be motivated by her belief that “promise-keeping is right,” Olivia may not be, because whereas Marsha is in both the belief-like and desire-like state of this belief, Olivia is in the belief-like state only.

Smith’s argument certainly seems to give strong reason to drop anti-Humeanism. If an agent is motivated by her moral belief that “ $\phi$ -ing is right” *only if* this belief has both directions of fit, and, this belief has both directions of fit *only if* it has two distinct contents, one of which is belief-like and the other desire-like, then we have no choice but to conclude that a desire, or some desire-like state, must always be present for motivation. But this is simply the claim of the HTM; hence, Smith’s argument not only supports a denial of anti-Humeanism, but an acceptance of the HTM as well.

To be clear, if the version of anti-Humeanism we have been considering were true, or could be redesigned to fix the problems we covered, it would present a

---

<sup>120</sup> Ibid.

<sup>121</sup> It seems that Smith’s ultimate suggestion here is that moral beliefs are better thought of as belief-desire composites. Smith does not explicitly endorse this view, but it is reasonable to think he does in light of something he says about other states that apparently have both directions of fit, e.g. wishes. Smith says, “...wishing that p is just desiring that p and believing that p is very unlikely...” Ibid., 117.

devastating objection to Williams' argument. Williams claims that reasons, to be reasons, must be able to explain actions, and to do this, reasons must be related to desires. But, if moral beliefs alone could do the same work as desires, then Williams would be unable to forge the link he wants between reasons and desires. For the time being, though, I think Smith's argument demonstrates that moral beliefs alone cannot do the same work as desires. Thus, unless we can be given some other account of motivation that entirely excludes desires I think premise 2 of Williams' argument is secure. Since, as I have mentioned, premise 3 requires no support of its own, but must be granted as a definitional point, this concludes my case for the soundness of Williams' argument. The next chapter will be devoted to more general problems facing Williams' conclusion.



### CHAPTER 3: THE WIDER ISSUES OF WILLIAMS' ARGUMENT

The previous two chapters were devoted to the validity and soundness of Williams' argument against external reasons. In short, I argued that Williams' argument is not only valid but sound as well. The purpose of this chapter is to take a step back and consider the wider problems facing Williams' conclusion. In the first section, I want to return to the issue of the potential consequences Williams' conclusion has for the moral domain.<sup>122</sup> Since moral reasons for action are often treated as external in Williams' sense, one might object that Williams' conclusion leaves us with an inadequate view of moral reasons for action. This is a legitimate concern, and I think it creates a burden for Williams to explain, on his internalist view, how we might get along without external reasons. I will present what I take to be Williams' most compelling response to this challenge. In sections two and three, I will confront two more objections which, somewhat like the first, allege that Williams' conclusion has certain unacceptable consequences. The first of these claims that Williams' conclusion leaves us with a purely instrumental theory of practical reason. This is problematic for some, because practical reason seems to be more than merely means-end calculation. The second claims that Williams' conclusion implies a form of egoism, a view which is, in any

---

<sup>122</sup> See chapter 1, section 3.

form, commonly thought to be deficient in several ways. If Williams' conclusion really is tied to either of these views, I think it would require us to reconsider the premises of Williams' argument more carefully. However, I will argue that neither of these views is a necessary consequence of Williams' conclusion, and therefore that the problems of these views cannot be held against Williams' argument. In the final section of this chapter, I want to look at two ways the critic of Williams' argument might neutralize the impact of his conclusion. Part of the sting of Williams' conclusion, it seems to me, is that without external reasons, there seems to be no basis for positing universal reasons, i.e. reasons for action shared by everyone. But, one might object, even if we grant Williams' conclusion, we need not admit that there are no universal reasons, because there might yet be universally shared *internal* reasons. I will examine two ways one might plausibly support such a possibility. Since Williams has little to say on this matter I think we must conclude, at the very least, that the possibility of universal reasons is an open issue for Williams' view.

### **§1. The Burden of Explanation**

Williams is somewhat unapologetic about his conclusion that there are no external reasons for action. While he admits that "People do say things that ask to be taken in the external interpretation," that is, that people commonly make external reason claims about themselves and others, Williams thinks, "all [such] external reason statements are false."<sup>123</sup> For Williams, people are simply in error about the fact that there are any external reasons. Of course, this is going to come as a surprise to anyone

---

<sup>123</sup> Williams, *Moral Luck*, 106-109.



who is sympathetic to, what I referred to previously as, *reasons externalism*.<sup>124</sup> In essence, reasons externalism is the view that an agent's reasons for action need not depend on his or her desires, i.e. that at least some reasons for action are external. To return to our original example, when a husband is self admittedly insensitive to his wife, and we claim that he has a reason to treat her better, and he replies "I don't care. Don't you understand? I really do not care," the reasons externalist might insist that, despite the husband's lack of desire, he nevertheless has a reason to treat his wife better.<sup>125</sup>

For many, the externalist stance here is going to seem undeniable – "of course the man has a reason to treat his wife better; what does it matter what *he* wants?" According to Williams, however, if there truly is no element in the husband's existing *S*, or a sound deliberate route that leads to such an element, which gives him a reason to treat his wife better, then there really is no reason for him to do so. Since Williams' conclusion so plainly runs against reasons externalism, a common and strongly held view, I think Williams picks up a burden, as I have suggested, to explain to the reasons externalist how we can get by without external reasons and genuinely true external reason statements. In what follows I want to present what I take to be Williams' best attempt to carry this burden. It is important to note from the outset that, while Williams denies that external reason statements are ever true, because there are no external reasons, he does not mean to imply that external reason statements are unimportant, or even that we should give up such externalist discourse. For Williams, as

---

<sup>124</sup> See chapter 1, section 3.

<sup>125</sup> Williams, *Making Sense of Humanity*, 39.

we will see, external reason statements play a special role in our deliberations about what we ourselves and others have reason to do.

To begin with, Williams acknowledges that reason discourse is often messy, because, for one thing, people do not explicitly express themselves in terms of internal and external reasons. We would not explicitly say, for example, that the husband, from before, has an *external* reason to treat his wife better; rather, we would say more simply that he has *a* reason to do so. Moreover, for any given reason claim we make about someone, it is often the case that we have in mind more than just one reason. When we advise other people, and say things like “you have a good reason to get an education,” we typically mean for this to cover multiple reasons for action, and not merely one. As a result, Williams thinks, it is difficult to identify and isolate exact instances where people truly intend for their reason statements to be taken singularly and externally, i.e. asserting the existence of exactly one external reason.<sup>126</sup> Nonetheless, Williams suggests, it must be the case that at least sometimes a “[reason] statement is...offered as standing definitely outside the agent’s *S*.”<sup>127</sup> In such cases, however, Williams argues that, since there are no external reasons, these reason statements must really be “something else misleadingly expressed.”<sup>128</sup> Williams’ initial

---

<sup>126</sup> Williams, *Moral Luck*, 111.

<sup>127</sup> *Ibid.*, 111.

<sup>128</sup> *Ibid.*, 111.



suggestion for this is that “those who use these words often seem, rather, to be entertaining an optimistic internal reason claim.”<sup>129</sup>

It is not entirely clear what Williams means by an “optimistic internal reason claim,” but we can take a clue from other things he says in the same passage. For Williams, external reason statements, like the claim that the insensitive husband has a reason to treat his wife better, might mean “little more than that things would be better if the agent so acted.”<sup>130</sup> Of course, the reasons externalist would likely want for this claim to mean something stronger than this, but, at the very least, I think this is something many people would be willing to accept – things *would be* better if the husband was nicer to his wife. Now, supposing the husband disagrees, because he cannot see why, from his point of view, things would be better if he did this, then, according to Williams, our reason statement really amounts to a “bluff.”<sup>131</sup> I think Williams’ idea here is something like this: when we openly declare to the man that he has a reason to treat his wife better, our declaration is literally false, but we express it with the hope that the husband will take it on and act accordingly.<sup>132</sup>

We might analogize what Williams has in mind here with another common practice to clarify what he means. It is often the case that when we see that a friend

---

<sup>129</sup> Ibid., 111.

<sup>130</sup> Ibid.

<sup>131</sup> Ibid.

<sup>132</sup> Elsewhere Williams clarifies this idea: “I suspect what are *taken for* external reason statements are often, in fact, optimistic internal reason statements: we launch them and hope that somewhere in the agent is some motivation that by some deliberative route might issue in the action we seek.” Williams, *Making Sense of Humanity*, 40.

lacks the courage or self-confidence required to overcome some problem he or she is facing, we nonetheless say things like “I know you can get through this, you are such a brave person.” In such cases, our statements might be literally false, but they are certainly not expressed without purpose. Sometimes a bluff like this from a friend is just enough to make a significant difference in someone’s confidence and behavior. In the same way, we might explain, an external reason bluff, simply because it is expressed by a friend, family member, or some other respected individual, might be enough for a person, who would typically deny the reason claim, accept it and alter his or her behavior.

By itself, the suggestion that external reason statements are really “bluffs” might not be enough to satisfy the reasons externalist. She might respond, “When I assert that the man has a reason to treat his wife better, I am not bluffing as I would in a card game – I genuinely believe that the man has a reason to change his conduct.” This is an entirely natural and reasonable response, though I think that Williams has a few more compelling things to say on the matter. In a later piece, Williams offered a more detailed description of the notion of an “optimistic internal reason claim.” He says:

...the claim that *A* has [an external] reason to  $\phi$  is not strictly true, by internalist standards, at the moment that it is made, but the very fact that it is made can help elicit a more general motivation from the agent’s *S*, such as the desire to have the respect of people like the speaker, and this motivation together with the recognition of those people’s desire or demand that he should  $\phi$  can indeed bring it about that he (now) has a reason to  $\phi$ .<sup>133</sup>

---

<sup>133</sup> Bernard Williams, “Postscript: Some Further Notes on Internal and External Reasons,” in *Varieties of Practical Reasoning*, ed. Elijah Millgrim (Cambridge, Massachusetts: MIT Press, 2001), 95.



It seems to me that this explanation takes the spirit of external reason claims more seriously. For Williams, when we express external reason claims about others, we are no longer simply hoping that an unconvinced individual will pick up the claim as before, but rather, we seem to be expressing a sincere concern or care that such an individual picks it up and acts accordingly. Complex motivations are now involved, like admiration and respect, and the act of uttering the claim itself is clearly meant to have direct effect on the individual's *S*. In this way, according to Williams, such reason claims, by the manner in which they are expressed, can help themselves to become true in the internalist sense.<sup>134</sup> So, to return to the insensitive husband, Williams would say that even though, as we may suppose, the reason claim about the husband, i.e. that he has a reason to treat his wife better, is literally false at the time it is made, the fact that the claim is expressed with urgency, care and concern by someone the husband loves, respects, or even fears, will likely produce in the man some desire or other which will make it true, in the internalist sense, that he has a reason to change.

At this point we can see that, on Williams' account, the non-existence of external reasons need not be a matter of serious concern. For Williams, the internalist position can offer everything the reasons externalist might want; internal reason statements can alter the behavior of people like the insensitive husband just as well as external reason statements could. Thus, Williams insists, we can get along just fine without external reasons and genuinely true external reason statements.

---

<sup>134</sup> The ideas here call to mind intervention settings where people come together with expressed purpose of changing a loved one's behavior, a person who, by internalist standards, might literally not have a reason to change until after the intervention.

I think Williams' explanation certainly helps to alleviate the externalist concern associated with his conclusion, though his explanation may not capture *everything* the reasons externalist might want. One likely objection would be this: on Williams' internalist position, we cannot ultimately criticize someone, in the way we want, for failing to pick up, and act in accordance with, one of these "optimistic internal reason claims." Imagine, for example, that even after sitting down with the insensitive husband for a good while, his respected peers cannot convince him that he has a reason to change his behavior with regard to his wife – that is to say, in Williams' terms, they cannot effect a substantial change in the husband's *S* to the extent that he himself acknowledges that he has a reason to change. In this case, the reasons externalist wants to say, regardless of the husband's stubbornness, he still has a reason to treat his wife better, and, if he fails to act on this reason, then he seems to be guilty of some kind rational failing.<sup>135</sup> On Williams' account, it could be urged, we simply cannot make this kind of criticism.<sup>136</sup>

It seems to me that this last objection undeniably hits its mark. I think Williams must concede that his conclusion strips away the possibility of certain kinds of criticisms the reasons externalist may want to make about people like the insensitive husband. On Williams' view, we simply cannot say of the insensitive husband, if he fails to pick up

---

<sup>135</sup> I tend to agree with the way Shafer-Landau describes such rational failings; he says "As I see it, to be rational is simply to be appropriately responsive in one's attitudes and behavior to sufficiently good reasons. To be irrational is to be insensitive, in one's attitudes and behavior, to sufficiently good reasons." Shafer-Landau, *Moral Realism*, 168.

<sup>136</sup> Shafer-Landau problematizes this point well: "To criticize someone, or someone's actions, as irrational is a quite strong kind of indictment. I think it important to preserve this feature in any analysis of (ir)rationality." Ibid., 168.



a given internal reason, that he is guilty of any sort of rational error. We may wonder, however, how deep this objection really cuts – why is it that we must secure this kind of criticism over any other? With regard to the insensitive husband, Williams fully admits that we cannot accuse him of any kind of rational failing, but he insists, we can say, for instance, that the husband is “...ungrateful, inconsiderate, hard, sexist, nasty, selfish, brutal, and many other disadvantageous things.”<sup>137</sup> Given this, Williams presses: “what is the difference supposed to be between saying that the agent has a reason to act more considerately, and saying one of the many other things we can say to people whose behavior does not accord with what we think it should be?”<sup>138</sup> For Williams, given that these types of criticisms are available to the reasons internalist, and that there seems to be no significant difference between these criticisms and the kind of criticism the reasons externalist wants to make, it follows that no serious problems arise for his conclusion that there are no external reasons for action.

In light of these considerations, I think Williams offers a strong response to the problem of how to get along without external reasons. Surely this is not to say that the reasons externalist at this point should be perfectly content with Williams’ conclusion. Indeed, it may be an impossible task to accommodate every feature of the externalist position. But, of course, this would be to ask too much of Williams from the outset. At bottom, Williams was obligated to explain why his conclusion does not leave the reasons externalist without any resources whatsoever. This much, it seems to me, has

---

<sup>137</sup> Williams, *Making Sense of Humanity*, 39.

<sup>138</sup> *Ibid.*, 39-40.

been done. For the time being, then, we can lay this issue aside and turn to other problems facing Williams' conclusion.

## **§2. The Problem of Instrumentalism**

In chapters 1 and 2, I confronted very specific objections to Williams' argument only. Given this, one might caution that even if Williams' argument is safe from these objections, it might yet be vulnerable to other more general objections. This is a reasonable concern, and I know of at least two such objections especially worthy of our consideration. The first, as previously mentioned, is that Williams' conclusion leaves us with a purely instrumental theory of practical reason. For those who think practical reason is not merely instrumental, this accusation may, if true, provide good reason to cast suspicion on the whole of Williams' argument. The second objection takes a similar form: Williams' argument implies a form of egoism and since any form of egoism is unacceptable, so too must be Williams' conclusion. I think Williams' conclusion is guilty of neither charge, and to demonstrate this I will draw directly from Williams' work itself. I will devote the present section to the first objection, and the following section to the second.

The first objection, that Williams' argument leaves us with a purely instrumental theory of practical reason, is somewhat ambiguous, because there is at least one critic, Brad Hooker, who argues that Williams' argument *presupposes* such a theory, and thus is question-begging.<sup>139</sup> Elijah Millgram, on the other hand, seems to suggest that

---

<sup>139</sup> Brad Hooker, "Williams' Argument against External Reasons," *Analysis*, 47, no. 1 (1987), 5-25.



Williams' argument *implies* a purely instrumentalist view.<sup>140</sup> I believe both of these objections are mistaken, but we need only silence Millgram's version to show this, because, it seems to me, if it is false that Williams' argument implies a certain view, then it must be false that it presupposes that view as well. But before we confront Millgram's objection, we need to get clear on what an instrumental theory of practical reason is, and why this is supposed to count against Williams' argument. For simplicity, I will follow Millgram and hereafter refer to such a view as *instrumentalism*.

According to Millgram, "Instrumentalism holds that all practical reasoning is means-end reasoning, that is, that figuring out what to do is entirely a matter of determining how to achieve one's goals or satisfy one's desires."<sup>141</sup> An example of instrumental reasoning, then, would be deciding to take the bus downtown, rather than biking, because doing so would be fastest way to get there. Of course, deciding to go downtown at all requires some desire, like the desire to meet a friend or to go shopping. Thus, taking the bus, in this case, represents the best means to achieving one's end to be downtown.

Instrumental reasoning like this, in itself, is not all that controversial. Everyone knows what it is like to figure out how to achieve some goal, and how to discover the best way to do so. The real point of contention is whether all practical reasoning is instrumental, that is, whether practical reasoning is always a matter of simply finding the means to one's ends. As it stands, this question is far too large and complex for us

---

<sup>140</sup> Elijah Millgram, "The Current State of Play" in *Varieties of Practical Reasoning*, ed. Elijah Millgram (Cambridge, Massachusetts: MIT Press, 2001), 4-5.

<sup>141</sup> *Ibid.*, 4.

to answer here. What we can do, however, is give at least some reason why someone might think practical reasoning must be more than merely instrumental. From here, we will be able to see why, supposing Williams' view implies a version of instrumentalism, someone might take Williams' conclusion to be unacceptable.

But first, why exactly would someone think that Williams' conclusion implies such a view? The line of thought behind this is simple: Williams claims that all reasons for action necessarily concern desires and their satisfaction. More specifically, as Williams' conclusion implies, an agent cannot have a reason to  $\phi$  unless she has some desire or other that will be served or furthered by her  $\phi$ -ing. From this it appears to follow that any reason for action an agent comes to as a result of practical deliberation will be one that identifies a certain course of action as a means to satisfying one of this agent's desires. But, in essence, this is the form of an instrumental theory of practical reason; hence, if Williams' conclusion is true, it seems that the only type of practical reasoning is means-end calculation. Now, why might this be a problem for Williams?

Millgram sets up the problem for instrumentalism nicely: "the instrumentalist position usually has it that practical justification bottoms out in the desires one just has: you can reason about to how to get what you want, but not about what to want in the first place."<sup>142</sup> This is problematic for many, because it seems to have the consequence that while we are able to criticize someone for not taking the best means to her ends, we are not able to criticize that person's ends themselves. Imagine, for example, a man whose sole desire in life is to collect pancakes. This man, we may suppose, spends all of

---

<sup>142</sup> Ibid., 4.



his time figuring out how best to get his hands on more pancakes, even at the cost of losing his friends and family. Now, the objector might argue, if instrumentalism is true, then we could only ever criticize this man for not taking the best means to achieving his end of collecting pancakes, e.g. for not buying the cheapest pancake batter, but we could never criticize the man's end itself. Yet, the objector might insist, it seems that we *can* criticize the man's end itself, e.g. because pancake collecting is not a worthwhile pursuit in itself; by *modus tollens*, instrumentalism must be false.

Certainly there are ways in which an instrumentalist might reasonably respond to this criticism. For our purposes, however, such a response would be irrelevant; we can simply grant that instrumentalism faces this problem. Furthermore, we can even grant that if Williams' argument implies a version of instrumentalism, his position faces this problem as well. My contention here is that Williams' argument does not imply a version of instrumentalism, and therefore that his position sidesteps such problems from the outset. To show this, we can draw from something Williams says in developing his version of reason internalism once he has drawn the conclusion that there are no external reasons. Recall that from chapter 1, reasons internalism is simply the view that all reasons for action are internal, and that Williams' particular brand of this view uses the notion of a sound deliberative route.<sup>143</sup> We characterized his position as follows:

**Williams' reasons internalism:** "A has a reason to  $\phi$  only if... [A] could reach the conclusion to  $\phi$  by a sound deliberative route from the motivations... [A] already has."<sup>144</sup>

---

<sup>143</sup> See chapter 1, section 2.

<sup>144</sup> Williams, *Making Sense of Humanity*, 35.

In contrast to this view is “basic reasons internalism,” which holds that an agent has a reason to  $\phi$  just in case  $\phi$ -ing will satisfy one of her desires. According to Williams, this basic version of reasons internalism “supposes that  $\phi$ -ing has to be related to some element in [an agent’s]  $S$  as a causal means to an end.”<sup>145</sup> Thus, Williams thinks, an overly simplistic version of reasons internalism implies a form of instrumentalism, because  $A$ ’s reason to  $\phi$  always depends on  $A$ ’s  $\phi$ -ing being a causal means to satisfying one of  $A$ ’s desires. But, for Williams, this does not mean that his version of reasons internalism implies the same. With regard to instrumental reasoning Williams says:

...there are much wider possibilities for deliberation, such as: thinking how the satisfaction of elements of  $S$  can be combined, e.g. by time-ordering; where there is some irresolvable conflict among the elements of  $S$ , considering which one attaches most weight to (which, importantly, does not imply that there is some one commodity of which they provide varying amounts); or, again, finding constitutive solutions, such as deciding what would make for an entertaining evening, granted one wants entertainment.<sup>146</sup>

This passage suggests that, for Williams, practical reasoning can be a very complex process involving much more than finding a means to an end, even – perhaps most importantly – attaching various weights to the ends themselves. Elsewhere Williams makes similar points about his position: “One of the most important things deliberation does, rather than thinking of means to a fixed end, is to think of another line of conduct altogether, as when someone succeeds in breaking out of a dilemma.”<sup>147</sup> This point connects back to something Williams’ builds into his notion of a sound

---

<sup>145</sup> Williams, *Moral Luck*, 104.

<sup>146</sup> *Ibid.*, 104.

<sup>147</sup> Williams, *Making Sense of Humanity*, 38.



deliberative route: the process of deliberation, through an exercise of the imagination “can create new possibilities and new desires.”<sup>148</sup> For Williams, then, it is not as if we are born with a predetermined and fixed set of desires, from which we can only determine how to satisfy them. The process of deliberation itself, according to Williams, can destroy old desires and produce new ones, which can then be ranked as more or less important. An agent may, for example, discover out of the blue that she likes to paint, and further that she wants more than anything to become an artist in life. For Williams, the reasoning involved here may, in part, be described as instrumental, e.g. the woman must figure out how she might obtain art supplies, but not all of it can be; the choice to become an artist at all, or deciding that painting is the most important thing to do over all else, is not a matter of pure means-end calculation. The only conclusion it seems we can draw from these considerations is that Williams not only avoids a pure instrumentalism about practical reasoning, but that he consciously does so.<sup>149</sup>

At this point, the critic may concede that Williams’ conclusion need not be tied to instrumentalism about practical reason, but she might yet wonder whether the view of practical reason Williams leaves us with is not deficient in some way. More specifically, one might question, can Williams’ view of practical reason really allow us to criticize people like the pancake collector from before? This is important for some,

---

<sup>148</sup> Williams, *Moral Luck*, 105.

<sup>149</sup> Illustrative of this last point is something Williams explicitly says after describing his view: “Can we [the reasons internalists] define notions of rationality which are not purely means-end? Yes.” *Ibid.*, 112.

because the pancake collector's activity obviously exhibits some sort of practical irrationality, and thus unless Williams' view can account for this intuition, his view, in addition to his conclusion, must remain suspect.

In response to this I think we can say, on Williams' behalf, that we can at least criticize the pancake collector for failing to take the best, e.g. the most convenient, economical, or pleasant, means to his own end, supposing he does.<sup>150</sup> Our criticism might aim to show, for instance, that given the man's desire for pancakes, he could have purchased more pancake batter if he had shopped at one store over another. But, of course, this is no more than a criticism of the man's means, a criticism any instrumental theory of practical reason could make. The real question, one might press, is whether, on Williams' view, we can criticize the pancake collector's end itself, something an instrumental theory seems to be incapable of doing. Initially, I think this poses no special problem for Williams' view. We might say with Williams, for instance, that the pancake collector has failed to recognize the fact that he has other more important ends, e.g. leading a productive life in society, which may ultimately be undermined by collecting pancakes all the time. Or perhaps, we might say with Williams, the man has failed to exercise his imagination sufficiently, so as to uncover sound deliberative routes

---

<sup>150</sup> We should note, however, that on Williams' view we might not, in certain extreme cases, ultimately be able to do this. This is because, for Williams, the desire "to take the correct means to his ends..." is itself a contingent element in the pancake collector's *S*. In spite of this, Williams reassures us, "In general, the aim of getting things right in such ways is part of any agent's interest as a rational deliberator, and the aim can be assumed to figure in any rational agent's *S*..." Thus, insofar as the pancake collector shares this common desire, i.e. roughly the desire to have one's desires satisfied, we can easily criticize him for failing to taking the best means to his own ends. Williams, "Postscript: Some Further Notes on Internal and External Reasons," 92.



to other ends he might find more worthwhile. These certainly seem to be ways, on Williams' view, that we can criticize the pancake collector's end itself.

One might interject here, however, that while these responses show how, on Williams' view, we can criticize the pancake collector's end itself, the problem is that these criticisms must always be made relative to other elements of the man's *S*, i.e. the man's desire may be criticized for undermining other more important desires he has, or for preventing the creation of new desires through processes of sound deliberation. But this leaves open the troubling possibilities that the pancake collector truly has no other more important desires, or that there truly are no sound deliberative routes to other more worthwhile ends. In such cases, it might be urged, Williams' view leaves us with no tools to criticize the pancake collector's end itself.

With regard to extreme cases like this, where a man's sole, most important desire is to collect pancakes, I think Williams must bite the bullet. Ultimately we cannot say, on Williams' view, that the pancake collector here is guilty of failing to act on better reasons, because for him there are none. Though, in defense of Williams, we might note that this objection is similar in form to the objection considered earlier this chapter involving the insensitive husband. In extreme cases, where an agent has certain unattractive elements in his *S*, e.g. the desire to treat his wife poorly, or the strong desire to collect pancakes, we may not be able to say of this agent that he has a reason to act otherwise, but, according to Williams, this does not mean that we cannot say other things about him, e.g. that he is unwise, foolish or inconsiderate. Thus, while we

may agree that both of these objections stand, I think we should not agree that either is all that devastating to Williams' view.

To summarize, Williams' conclusion does not leave us with a purely instrumental theory of practical reason. It is clear from Williams' work that, despite the apparent ties between his view and instrumentalism, Williams consciously aims to offer a more sophisticated account of practical reason. Furthermore, even though Williams' view may not ultimately allow us to criticize people in extreme cases with strange or unattractive desires, we should not allow this to count too heavily against his view, because, as Williams says, there seems to be no significant difference between this kind of criticism and others we might want to make about such people.

### §3. The Problem of Egoism

I want to turn now to the objection that Williams' conclusion implies a form of egoism. This objection grows out of the concern that, according to Williams' conclusion, an agent's reasons for action always only depend on what would serve or further *her* desires. Marsha's reasons for action, for instance, never directly depend on anyone's desires but her own. For some, this seems to suggest that Williams' conclusion supports an overly *egoistic* view of reasons for action in general; all Marsha's reasons for action involve the satisfaction of Marsha's desires.<sup>151</sup> I think this accusation, if true, would help to cast suspicion on Williams' argument as a whole.

---

<sup>151</sup> Initially, this objection may sound similar to the one posed by the reasons externalist concerning individuals like the insensitive husband, but I believe it constitutes a distinct challenge, since the point of this objection is not that, on Williams' view, particular individuals may not ultimately have particular reasons to act morally, but rather that *everyone's* reasons for action must *always* flow from egoistic or self-serving considerations.



As it stands, however, the accusation that Williams' conclusion implies a form of egoism is rather vague, since there are numerous forms of egoism available. In very general terms, we can distinguish between at least three different basic forms: psychological, ethical, and rational egoism. From the outset, however, I think we can exclude both psychological and ethical egoism from consideration, since Williams' conclusion could not possibly imply either of these views. Firstly, psychological egoism is an empirical view about the nature of human action, which maintains that each person always acts with only one thing in mind: the satisfaction of his or her own desires. Williams' conclusion, in contrast, is a conceptual claim about the conditions under which an agent could have reasons for action at all. Given this, the truth of Williams' conclusion is entirely consistent with the falsity of psychological egoism; for, we could easily imagine a world in which all reasons for action are internal, as Williams says, but that people do not always in fact act with the satisfaction of their own desires in mind. Secondly, ethical egoism is a prescriptive view, according to which an agent is morally obligated to perform an action just in case that action contributes to the satisfaction of her desires.<sup>152</sup> Williams' conclusion concerns reasons for action, not moral obligations, and as such makes no prescriptions about what agents morally ought to do. This leaves the possibility that Williams' conclusion implies some version of rational egoism, which seems to me to be the most plausible possibility from the beginning. To assess whether this is true, we should first pause to consider rational egoism itself.

---

<sup>152</sup> This is of course not the only way one might specify ethical egoism.

There are numerous ways one might spell out rational egoism. Given this, it will be helpful to draw upon a few ways philosophers have described the view. According to Shafer-Landau, a strong version of rational egoism holds that "...one has a reason to  $\phi$  if and only if  $\phi$ -ing will *best* promote one's self-interest."<sup>153</sup> So stated, this statement is not too far from Williams' conclusion, since unlike psychological and ethical egoism it at least concerns the conditions under which an agent has a reason for action. Yet, we can see quickly that this version of rational egoism could not possibly be implied by Williams' conclusion, since Williams' conclusion says nothing about what an agent has *best* reason to do. According to Williams,

...an internal reason statement does not apply only to that action which is the uniquely preferred result of the deliberation. 'A has a reason to  $\phi$ ' does not mean 'the action which A has overall, all-in, reason to do is  $\phi$ -ing.' He can have reason to do a lot of thing which he has other and stronger reasons not to do.<sup>154</sup>

However, rational egoism can easily be weakened to accord better with what Williams' says here. Shafer-Landau offers the following as a more subdued version of the view: "Rational egoism is the thesis that one has a reason to  $\phi$  if and only if  $\phi$ -ing will serve one's self-interest."<sup>155</sup> This version is certainly closer to Williams' view, but it remains somewhat unclear what Shafer-Landau means by "self-interest." At the very least, I think it is reasonable to say that, for Shafer-Landau, an agent's self-interest includes things like her meeting her wants and desires. Another version, from David Brink, says a little more on this point: "Rational egoism...says that an agent has reason

---

<sup>153</sup> Shafer-Landau does not endorse this view. Shafer-Landau, *Moral Realism*, 194.

<sup>154</sup> Williams, *Moral Luck*, 104.

<sup>155</sup> Shafer-Landau, *Moral Realism*, 194.



to do x just in case... x would contribute to *his own* interest, welfare, or happiness.”<sup>156</sup>

Given the inclusion of an agent’s welfare here, it seems to me that, for Brink, an agent’s self-interest includes not merely her wants and desires, but also things like her physical and psychological well-being. So, for example, we might say that eating healthy food contributes to Marsha’s self-interest even if she has no desire to eat such food. In this way, according to a rational egoist, an agent may have a reason to  $\phi$  even though she has no desires that will be served or furthered by  $\phi$ -ing, so long as  $\phi$ -ing at least contributes to her physical or psychological well-being.

On this interpretation of self-interest, however, I do not think that Williams’ conclusion could possibly imply rational egoism as so far described. This is because, on Williams’ view, an agent might be completely unmotivated to pursue what is in his best interest, physically or psychologically speaking.<sup>157</sup> Williams’ considers a man who needs to take a certain kind of medicine, but “consistently and persuasively denies any interest in preserving his health.”<sup>158</sup> While the rational egoist might insist that the man has a reason to take the medicine, Williams would not. With regard to cases like these, Williams says this:

If an agent really is uninterested in pursuing what he needs; and this is not the product of false belief; and he could not reach any such motive from motives he has by ...[a process of sound deliberation]; then I think we do have to say that in the internal sense he indeed has no reason to pursue these things.<sup>159</sup>

---

<sup>156</sup> Like Shafer-Landau, Brink describes but does not support this view. Brink, *Moral Realism and the Foundations of Ethics*, 67.

<sup>157</sup> Williams, *Moral Luck*, 105.

<sup>158</sup> *Ibid.*, 106.

This does not mean, however, that Williams' conclusion does not imply a weaker version of rational egoism. Both Shafer-Landau and Brink discuss one prominent type of rational egoism which restricts the concept of an agent's self-interest to an agent's desires.<sup>160</sup> On this version of rational egoism, what Shafer-Landau and Brink call the "desire-satisfaction" view, the notion of an agent's welfare is dropped so that only desires remain. With some liberty, we may characterize this new version of rational egoism in the following way: an agent has a reason to  $\phi$  if and only if  $\phi$ -ing will contribute to the satisfaction her desires. This characterization of rational egoism is undeniably closer to Williams' view, and although it is weaker, it may nonetheless appear as unpalatable as the stronger version we began with. According to this version, reasons still seem to be egoistic and self-serving, because an agent's reasons for action are always a function of what will contribute to the satisfaction of *this* agent's desires only, e.g. Marsha has a reason to  $\phi$  when and only when  $\phi$ -ing serves her *own* personal ends.

Now, since, according to Williams, no reasons for action are external, it follows that all reasons for action depend, in some way or other, on the satisfaction of an agent's desires. Given this, we might wonder how this latest version of rational egoism is *not* implied by Williams' conclusion – both concern an agent's reasons for action and the satisfaction of an agent's desires. The first thing we should say is that, strictly speaking, this version of rational egoism cannot fall out of Williams' conclusion, since it

---

<sup>159</sup> Ibid., 105.

<sup>160</sup> Shafer-Landau, *Moral Realism*, 194; Brink, *Moral Realism and the Foundations of Ethics*, 67.



asserts both a necessary and sufficient condition for having a reason for action, whereas Williams' conclusion implies only a necessary condition. But suppose we drop the sufficient condition from our current version of rational egoism – does Williams' conclusion then imply the view? I think so – both views now seem to state essentially the same thing: an agent has a reason to  $\phi$  only if  $\phi$ -ing will serve or further (i.e. contribute to the satisfaction of) this agent's desires.<sup>161</sup> Is this a problem for Williams' view? That is, do the problems associated with rational egoism transfer to Williams' conclusion?

I think not. This is because Williams does not place any specific constraints on the content of an agent's desires. The concern with rational egoism is that seems to endorse egoistic or self-serving actions and behaviors. Williams' conclusion need not be placed in the same boat. With regard to the elements of an agent's *S*, i.e. the complete collection of an agent's desires, Williams says: "Above all, there is of course no supposition that the desires or projects of an agent have to be egoistic; he will, one hopes, have non-egoistic projects of various kinds, and these equally can provide internal reasons for action."<sup>162</sup> This clarification, it seems to me, helps to get at the heart of why someone might think Williams' view leads to an egoistic doctrine. When Williams says that reasons depend on desires, this may make one think that the reasons must somehow be selfish in nature. This is of course not true. On Williams' account, an agent may have reasons to perform all kinds of unselfish acts; an agent may even have

---

<sup>161</sup> Of course, Williams' view is slightly more sophisticated, since it includes possible desires, i.e. ones arrived at through a process of sound deliberation.

<sup>162</sup> Williams, *Moral Luck*, 105.

reason to give up her own life, provided she desired the outcome of that act more than continuing to live.<sup>163</sup> In light of these considerations, I think we must conclude that although Williams' conclusion implies a (weak) version of rational egoism, this is not problematic, because desires, on Williams' account, need not be selfish or self-serving in nature.

From here, we can see that Williams' argument is safe from the more general objections that if we accept the conclusion, we must be committed to instrumentalism or some form of egoism (or both). Williams' argument leaves us with neither view, and thus the faults of these views cannot be used against Williams' argument.

#### **§4. Universal Reasons**

The final challenge I want to consider is unlike all the others thus far considered, since it begins by granting the soundness of Williams' argument, and therefore the truth of its conclusion. The claim that there are no external reasons is unacceptable for some, in part, because it seems to suggest that there are no reasons for action shared by all agents as such, i.e. that there are no *universal* reasons for action. The line of thought behind this is simple: supposing, as Williams says, that reasons depend on desires, and that there is, empirically speaking, a wide divergence of desires between and even within different cultures, it appears to follow that no reasons for action are universally shared by all people, i.e. there are no universal reasons for action. This is supposed to be problematic, because it seems to lead to *relativism* about practical reason: an agent's

---

<sup>163</sup> Williams says, "There is no contradiction in the idea of a man's dying for a ground project – quite the reverse, since if death really is necessary for the project, then to live would be to live with it unsatisfied, something which, if it really is his ground project, he has no reason to do." Ibid., 13.



reasons for action are always relative to her own concerns even if they conflict with the concerns of others. Despite the fact that this is only part of the reason why some take issue with Williams' conclusion, it seems to me that this idea in fact represents one of the intended upshots of Williams' conclusion. Of course, the notion that reasons must be tied to desires is central to Williams' argument, but I think Williams wants use this idea, at least partially, as a stepping stone to help establish the more general claim that there are no universal, predetermined or fixed projects in life.<sup>164</sup>

Now, if it can be shown that there could be universally shared reasons even if Williams is right that all reasons for action are internal, then, I think, this would help to neutralize the impact of Williams' conclusion, because it would show that the non-existence of external reasons does not imply relativism about practical reason. Of course, on the assumption that Williams wants the further conclusion that there are no universally shared reasons for action, then the success of this challenge would also count as a blow to Williams' larger project. In way of preview, I think that Williams does not give us the resources to block this challenge, thus for good or bad, the possibility of universal reasons must remain an open issue for Williams' view.

---

<sup>164</sup> In fact, much of Williams' work seems to be devoted to this task. Evidence of this may be found in Williams' critiques of the "universal" and "impersonal" character of both Kantian and Utilitarian moral theories. For Williams, such theories sometimes demand that an agent forgo his own "project(s)" in the name of some moral course of action. This can be problematic, according to Williams, because "...[an agent's] present projects are the condition of...[his] existence, in the sense that unless...[he is] propelled forward by the conatus of desire, project, or interest, it is unclear why...[he] should go on at all..." In other words, for Williams, if an agent's projects are the only reason he lives at all, then anything that stifles these projects, threatens the very meaning of this man's life. Williams says "There can come a point at which it is quite unreasonable for a man to give up, in the name of the impartial good ordering of the world of moral agents, something which is a condition of his having any interest in being around in that world at all." Ibid., 12-14.

I would like to present two ways one might formulate this challenge. Both begin by accepting Williams' conclusion, i.e. that there are no external reasons for action, but then move on to challenge the further conclusion, i.e. that no universal reasons for action. The first is based on the idea that there might be "deep" psychological facts shared by all people, and the second exploits the vagueness of Williams' concept of practical deliberation. To begin with, let's take a quick look at the basic argument for the idea that Williams' position implies the non-existence of universal reasons, where a universal reason is one that is shared by all people as such:

(1) Reasons, in one way or another, always depend on desires.

(2) Different people have different desires.

(3) Different people have different reasons, i.e. there are no universal reasons.

For the sake of argument we should grant premise 1 here, because we want to see whether the non-existence of universal reasons follows from the claim that there are no external reasons for action. The first challenge aims to demonstrate the possibility of universal reasons by rejecting premise 2 of this argument, i.e. the claim that different people have different desires. This strategy may seem implausible at first, given the near undeniable divergence of peoples' desires, yet, I think it can be built up in a fairly compelling manner. We should note that anyone who posed this challenge would likely grant that many of our desires are not shared, because we are, for instance, brought up in different places, with different families, and different cultures. This much



is not disputed. The real point to be made is that perhaps there are at least some desires we do share in virtue of “deep” or widely shared psychological traits. This kind of suggestion is made by David Brink, a critic of Williams’ view, though not explicitly about reasons. In the following passage, Brink is concerned with obligations, but I think the point could be made about reasons as well:

If, for example, sympathy is, as Hume held, a deeply seated and widely shared psychological trait, then as a matter of contingent (but “deep”) psychological fact, the vast majority of people will have at least *some* desire to comply with what they perceive to be their moral obligations...<sup>165</sup>

The basic idea behind this challenge, then, is that since there are certain desires which are shared by nearly every person, due to widely shared psychological traits, we have strong reason to think that at least some reasons for action are universal. In this way, it can be argued, even if all reasons for action are internal, it might yet be the case that at least some reasons for action are shared by all people.

This challenge is very interesting, but I think it faces a few problems of its own. For one, even if we did discover such deep-seated psychological traits that ground certain desires, it might be argued that, given the inevitable variation among people, there will be at least some who do not share the desires required to participate in the alleged universal reasons, in which case, strictly speaking, they would not be *universal* reasons after all. I suspect, however, that the critic, like Brink, would ultimately tolerate this conclusion. Perhaps, he might respond, while it may be case that there are no reasons shared by *every* single person, it may nonetheless be the case that there are

---

<sup>165</sup> Brink, *Moral Realism and the Foundations of Ethics*, 49.

some reasons shared by the *majority* of people. This is a reasonable response, but there may be a stronger way to argue for the possibility of universal internal reasons. After all, this challenge ultimately depends on an empirical claim, which means that, at best, universal reasons could only be said to exist contingently. The next challenge is an improvement in this manner, since it aims to show that universal reasons might exist necessarily. Ironically, Williams introduces the possibility of such a challenge:

If it were true that the structure of practical reason yielded reasons of a certain kind as binding on every rational agent, then it would be true of every rational agent that there was a sound deliberative route from his or her *S* to actions required by such reasons.<sup>166</sup>

In other words, according to Williams, perhaps the nature of practical deliberation itself, as opposed to contingent psychological facts, conditions the possibility of universal reasons.

Thus far, we have discussed Williams' conception of practical deliberation in some detail, but we have yet to examine it critically. One prominent type of objection to Williams' view claims that he leaves his conception of practical deliberation, and, more specifically, his notion of a sound deliberative route, incredibly vague.<sup>167</sup> This is problematic, critics suggest, because without more specific constraints, the possibility is opened that anyone, no matter the contents of his or her existing motivational set, could be lead through some path of sound reasoning to any reason for action. For all Williams has said, one might argue, the insensitive husband, for instance, might have an

---

<sup>166</sup> Williams, "Postscript: Some Further Notes on Internal and External Reasons," 94.

<sup>167</sup> Korsgaard, *Skepticism about Practical Reason*; Cohon, "Are External Reasons Impossible?"; McDowell, "Might There Be External Reasons?"



internal reason to treat his wife better after all, since there is some undiscovered path of sound reasoning that would lead him to this conclusion. This is a very complex issue, and I am not certain that Williams does much to help his own view out. In fact, Williams appears to, not only admit the charge of vagueness, but even embrace it. In the earliest piece on the matter, Williams says:

But here it may be objected that [my] account of deliberation is very vague, and has for instance allowed the use of the imagination to extend or restrict the contents of the agent's *S*. But if that is so, then it is unclear what the limits are to what an agent might arrive at by rational deliberation from his existing *S*.<sup>168</sup>

And, immediately following this, Williams says:

It is unclear, and I regard it as a basically desirable feature of a theory of practical reasoning that it should preserve and account for that unclarity. There is an essential indeterminacy in what can be counted a rational deliberative process. Practical reasoning is a heuristic process, and an imaginative one, and there are no fixed boundaries on the continuum from rational thought to inspiration and conversion.<sup>169</sup>

It is difficult to know what to say for Williams here. On the one hand, Williams wants to rule out the possibility of reasons for action that apply to every person, but on the other hand, Williams wants to leave open his concept of practical deliberation to capture the way people really reason about what to do, even at the cost of setting his view up for a strong objection. Of course, we should recognize that quite a bit of effort would have to be spent demonstrating how exactly the nature of practical deliberation, in Williams' sense, could be such that any agent, no matter his or her existing desires, would be lead to the same reasons for action, provided he or she reasoned correctly. But given that

---

<sup>168</sup> Williams, *Moral Luck*, 110.

<sup>169</sup> *Ibid.*, 110.

Williams offers no reason to think that this could not be done in principle, at the very most, I think Williams may be allowed to keep his conclusion that there are no external reasons for action, but not the further conclusion that there are no universal reasons for action. This concludes my discussion of the wider issues facing Williams' argument. I will end by summarizing the major results of this thesis.

## **§5. Conclusion**

In this thesis I have attempted to clarify and critically examine Bernard Williams' argument against external reasons. In chapter 1, I set out to formalize and explain in some detail the steps of this argument. I view this as an important task in itself for two main reasons. First, although some critics and defenders have put forward various informal characterizations of Williams' argument, there has been little effort to reconstruct the argument carefully in premise and conclusion form. I suspect that this is the root of several of the misguided objections we have considered. Shafer-Landau's objection, for instance, that Williams' argument conflates two different kinds of reasons, arises only because of an inaccurate reading of premise 1.<sup>170</sup> Second, a more careful look at an argument can often help to uncover important assumptions that have previously gone unnoticed. To my knowledge, no one has, for instance, identified the interesting relationship between premise 1 of Williams' argument and the claim I have referred to as (R), i.e. *R* is a reason for *A* to  $\phi$  only if *A* is able to  $\phi$  at all. This is a crucial

---

<sup>170</sup> Additionally, I think Cohon's objection to premise 2, i.e. the various counterexamples designed to show that explanations of intentional actions need not reference an agent's desires, depends on an inaccurate reading of premise 2. An intentional act, in Williams' sense, is always one that is performed for a reason.



consideration for Williams' argument, because, as we have seen, there are objections to premise 1 that presuppose that (R) is false without good reason.

In addition to formalizing Williams' argument in chapter 1, I defended its validity. But this was only the first level of evaluation. In chapter 2, I set out to discuss and defend the soundness of Williams' argument. Although I defended Williams' argument at nearly every turn, my primary aim was not to provide an airtight defense of each premise. Rather, my goal was simply to show that Williams' argument is stronger than some critics might suppose. Aside from the definitional claim of premise 3, I believe that premise 1 is the most difficult piece of Williams' argument to deny: all reasons for action must have the potential to explain an action performed for that reason. Many philosophers, including some of Williams' critics, are willing to grant this premise from the outset. Thus, if there is a weakness of Williams' argument, I think it must lie with premise 2: reasons cannot explain intentional actions unless they are related to an agent's desires. As we have seen, Williams assumes some version of the Humean theory of motivation, i.e. the view that beliefs are not sufficient, and desires are necessary, for motivation, to support this premise. I considered one prominent alternative to this theory, which I referred to as *anti-Humeanism*. If, as this view holds, moral beliefs could motivate agents to act all by themselves, then premise 2 would be undeniably false, because it would be possible to explain an agent's intentional actions without recourse to her desires. I drew from Smith's work, however, to show that this view seems to face serious problems of its own. I concluded that we have better reason

to stick with Williams on the issue of motivation, and therefore agree that premise 2 is secure for the time being.

In this final chapter, I have tried to supplement my critical examination of Williams' argument with a discussion of some of the wider issues facing it. Given that Williams' argument appears to have multiple unpalatable consequences, I felt it was important to show that Williams' argument need not be as ruinous as one might suppose. Even without external reasons, I think that Williams provides us with the resources to redirect the recalcitrant behavior of individuals like the insensitive husband. In addition to this, I argued that Williams' conclusion need not be burdened with the faults of instrumentalism or egoism. On these counts, at least, I believe Williams' argument has been fully exonerated. However, with regard to the final consideration, i.e. the possibility of universal internal reasons, I think Williams' conclusion is insufficient by itself to issue a decision either way; the non-existence of universal reasons does not necessarily follow from the non-existence of external reasons. Much more work would have to be done to rule out, or demonstrate, the possibility of such reasons for action. Thus, insofar as Williams, or his defenders, wants the further conclusion that there are no universal reasons for action, additional arguments are required. Independently of this further concern, however, I think that Williams' argument, along with its immediate conclusion that there are no external reasons for action, stands strong.