

University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

2011

Bernard Williams on Coercion and Reasons for Action

Joseph Stephenson

The University of Montana

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Stephenson, Joseph, "Bernard Williams on Coercion and Reasons for Action" (2011). *Graduate Student Theses, Dissertations, & Professional Papers*. 91.

<https://scholarworks.umt.edu/etd/91>

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

BERNARD WILLIAMS ON COERCION AND REASONS FOR ACTION

By

JOSEPH B. STEPHENSON

Bachelor of Arts, University of Maine at Farmington, Farmington ME, 2006

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Arts
in Philosophy

The University of Montana
Missoula, MT

Official Graduation Date: December 2011

Approved by:

Sandy Ross, Associate Dean of the Graduate School
Graduate School

Bridget Clarke, Associate Professor
Philosophy

Albert Borgmann, Regents Professor
Philosophy

Bari Burke, Professor of Law
The Law School

ABSTRACT

Stephenson, Joseph, Master of Arts, 2011

Philosophy

Bernard Williams on Coercion and Reasons for Action

Chairperson: Bridget Clarke

Understanding the nature of reasons for action and, specifically, how one can come to understand and criticize another's reasons for action can help us to better understand what possible outcomes we can hope for when engaging another agent. Understanding reasons for action may even give us some insight into what shared ground ethical theory can offer.

Bernard Williams' treatment of reasons for action informs much of the contemporary discussion. Williams argues that there are generally two classes of statements about reasons: internal and external. An internal reason, according to Williams, is a reason which depends entirely on an agent's motivations and perspective. Something about who the agent is and how she sees the world determines what internal reasons she has. An external reason is a reason to act no matter what motivations an agent might have or what her perspective.

Williams focuses much of his thoughts on external reasons around disclosing their unintelligibility. The problem which I deal with is that through the course of trying to disclose external reasons' unintelligibility, Williams claims that appealing to external reasons holds some inherent danger of coercion. Williams believes that if an agent who takes the role of advisor wants to convince another agent that an external reason applies to him, and rational persuasion fails, then he *has* to use coercion to persuade the agent to act.

I argue that Williams' claim about the relationship between external reasons and coercion is false. I show that he erroneously tangles external reasons with what he calls "the morality system", making external reasons appear necessarily coercive. By separating external reasons from the morality system, I

expose why external reasons need not entail any more coercive potential than Williams' version of internal reasons.

Introduction

All animals act for a reason. The lioness kills for food. The lion cubs play to hone their hunting skills necessary to thrive. The human has the unique capacity to reflect on the reasons why she acts, thus refining or rejecting various inclinations and making a reason, in a conceptual sense, her own. And so it is that whereas there are reasons why every animal acts, only a human can be said to *have* a reason to act. Humans bear unique responsibility for their reasons.

Human reasons for action can conceivably come from many places: desires, inclinations, instincts, traditions, ethics, biological necessity and so on. We care about human reason for action, aside from the more abstract philosophical concerns, for the sake of ethics. Understanding the nature of reasons for action and, specifically, how one can come to understand and criticize another's reasons for action can help us to better understand what possible outcomes we can hope for when engaging another agent. Understanding reasons for action may even give us some insight into what shared ground ethical theory can offer.

Bernard Williams' treatment of reasons for action informs much of the contemporary discussion. He focuses his thoughts on reasons around an agent's perspective or what he calls an agent's "motivational set," and most often refers to as an agent's "S." For Williams, S captures the way an agent experiences the world and the way the world presents itself to that agent. An S can include "dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects as they may be abstractly called."¹ Williams uses S to portray the contingency of human existence. An agent's S necessarily conditions the way she sees the world. Every agent gets thrown into the world with a certain set of attributes and has to deal with a certain scope of circumstances which limit her possible methods for navigating her surroundings. There is certainly no universal S, although many shared attributes most

¹ Bernard Williams "Internal and External Reasons", in *Moral Luck*. (Cambridge University Press: Cambridge, 1981) pg. 110-111.

likely exist among humans, and particularly among individuals who share the same culture. Furthermore, an agent's S is not statically given; it can change through new experiences. Deeper self or world understanding can add elements to or subtract elements from an agent's S.

Williams argues that there are generally two classes of statements about reasons: internal and external. An internal reason, according to Williams, is a reason which depends entirely on an agent's S. Something about who the agent is and how she sees the world determines what internal reasons she has. An external reason, for Williams, does not depend, in any way, on an agent's S. An external reason is a reason to act no matter what motivations an agent might have or what her perspective. Williams uses the example of Owen Wingrave to illustrate his view of an external reason.

Owen's father urges on him the necessity and importance of his joining the army, since all his male ancestors were soldiers and family pride requires him to do the same. Owen Wingrave 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. His father might have expressed himself by saying that *there was a reason for Owen to join the army*. Knowing that there was nothing in Owen's S which would lead, through deliberative reasoning, to his doing this would not make him withdraw the claim or admit that he made it under a misapprehension. He means it in an external sense.²

Owen's father does not appeal to anything in Owen's motivations or perspective in this example. The reason exists entirely separately from these considerations. Owen's father can apply the reason to Owen with no knowledge of Owen or requirement even for Owen to understand the reason.

Williams focuses much of his thoughts on external reasons around disclosing their unintelligibility. Very roughly, Williams argues that if nothing about who an agent is might lead her to act on, or even understand a reason, then she cannot be said to have a reason. On the other hand, if something about who the agent is may lead her to act on a reason, or she has a conscious motivation to act on that reason, then she simply has an internal reason. Williams believes that if an advisor says of an agent that

²Williams, 1981, pg. 105.

she has an external reason to act, the advisor must be mistaken or referring, perhaps unknowingly, to something else.³

Also, according to Williams, appealing to external reasons holds some inherent danger of coercion. In a rarely cited piece, which Williams wrote late in his career, he claims there is a necessary connection between appealing to external reasons and coercing an agent into action. “[T]he failings of an externalist account come out all the more clearly when we reflect on the kind of discussion that might be needed to convince an agent that [an external reason] applied to him, and how that discussion could hardly fail to be coercive.”⁴ Williams believes that if an agent who takes the role of advisor⁵ wants to convince another agent that an external reason applies to him, and rational persuasion fails, then he *has* to use coercion to persuade the agent to act.

Rather than arguing about whether external reasons exist or not, I will deal with a more tangible question: if external reasons did exist, would they, as Williams argues, have a necessary, or even a strong connection with coercion? I will argue that Williams’ claim about the relationship between external reasons and coercion is false. I will show that he erroneously tangles external reasons with what he calls “the morality system”, making external reasons appear necessarily coercive. By separating external reasons from the morality system, I expose why external reasons need not entail any more coercive potential than Williams’ version of internal reasons.

I devote Chapter I to elucidating Williams’ claim about external reasons and coercion. First, I draw out what Williams means by coercion. Then I sketch two extreme theories of reasons for action, the sub-Humean model of reasons and Williams’ conception of external reasons (both of which Williams tries to avoid) and explain how each extreme relates to coercion.

³ Williams, 1981 pg. 106-110.

⁴ Bernard Williams, “Values, Reason and The Theory of Persuasion”, in A.W. Moore ed. *Philosophy as a Humanistic Discipline*. (Princeton University Press: Princeton, 2008). pg. 118.

⁵ I will be using the term “advisor” to refer to the agent attempting to get another agent to act.

In Chapter II, I give an in depth discussion of Williams' own internalism, which he intended as a middle ground between the two extreme theories of reasons for action. I argue that Williams' most viable position for creating a middle ground fails because it does not offer significant protection against coercion as compared to a reasonable conception of external reasons.

Finally, in Chapter III, I can expose why Williams' worry about coercion and external reasons is unwarranted. I show why external reasons need not be conceived of as part of the morality system. Through the course of disembedding external reasons from the morality system, I will pinpoint just what makes a reason external, and give the rough contours of an alternative to Williams' conception of external reasons and how they have to factor into deliberation. The result will then fall out that, contrary to Williams' claims, there is no strong connection between external reasons and coercion as compared to the connection between his internal reasons theory and coercion.

Chapter I: Making Sense of Coercion in External Reasons

Normativity and Coercion

Although Williams maintains that external reasons can too easily incite coercion, he does see certain benefits in external reasons. The benefits Williams acknowledges in external reasons are that they can be used to criticize or advise another agent. External reasons act as grounds for criticism, and create an opportunity for shared deliberation.

These positive attributes of external reasons constitute what philosophers, including Williams, often call *normativity*. Normativity is an extremely contentious term among philosophers, but I will use it, as Williams does, to refer to how an agent may be wrong about her reasons in a way that warrants

correction.⁶ Reasons are normative if they can be used to say what an agent ought to or ought not to do, or have done.

Williams thinks of his internalism as a middle ground between a simplistic desire-based model of reasons for action and the model which supports external reasons as the staunchly objective entities he believes they have to be. He wants to push internal reasons as close to external reasons as possible, still keeping all reasons subject to an agent's perspective, but adding to them some normative force. This move allows internal reasons to be world guided *and* action guiding.

Externalism holds certain dangers, according to Williams, because if an advisor takes it that reasons have very strong and easily understood normativity, she may be able to ignore the agent's S and use external reasons to justify coercing the unresponsive agent into action. Before detailing the two extreme formulations of reasons for action, then, I will draw out Williams' definition of coercion. I can then disclose where Williams goes wrong in tying external reasons so strongly to coercion. I will suggest that what Williams means by coercion is not far from the ordinary usage of the term.

Williams gives no explicit definition of coercion, but in "Values, Reasons and The Theory of Persuasion" he gives us some insight into what he might mean by coercion through some constraints he puts on good deliberative assistance. Failing to fulfill the requirements for good deliberative assistance does not necessarily mean an advisor will use coercion, but I can still glean some insight into what Williams has in mind and add a point or two from some unrelated comments in "How Free Does the Will Need to Be?" where Williams gestures at coercion in a discussion of constraint. Good deliberative assistance requires that:

- (i) The assistant [or advisor] will be truthful, in the sense both of telling the truth and of helping the agent to discover the truth. An application of this is that the assistant will be truthful about his own procedures and motives, with the result that these can be transparent to the agent; the assistant has no hidden agenda in his dealings with the agent.

⁶ Bernard Williams, "Postscript" in Elijah Millgram ed. *Varieties of Practical Reasoning*. (MIT Press: Cambridge, 2001). pg. 92-93.
Stephenson

- (ii) [a] The assistant will try to make the best sense of the agent's S, and, in particular, [b] if there is a conflict between the assistant's and the agent's interpretations of the agent's S, the assistant will have some suitable explanation of the agent's misinterpretation. (This condition is not circular: it does not say that the assistant has to make the best sense of what the agent's reasons are.)⁷

If an advisor is honest with the agent, attempts to understand her S and is willing to talk through any misunderstandings, it is difficult to imagine how this conversation could possibly be coercive. Although not the only way for coercive-free persuasion, this does offer protection from coercion.

Coercion requires at least two actors, whether governments, individual agents or some combination. Generally an advisor coerces an agent into action when the advisor forces the agent to do something which she "did not originally want to do, or which [she] would not want to do unless the circumstances were exceptionally disadvantageous..."⁸ The fact that (i) requires the advisor to honestly disclose her motives and view of the world suggests that trickery can be coercive. The fact that (ii) requires that the advisor discusses and engages the agent's S suggests that if an advisor disregards an agent's own interests in prescribing action, the advisor may fall into coercion. Further, the fact that Williams wants to defer to the agent's own interpretation of her S suggests that the crucial requirement is that the agent be able to come to the conclusion to act on her own, voluntarily. For Williams, then, coercion seems to be when one agent tries to get another to act against her own wishes, whether through force, deception, trickery, irresistible incentive⁹ etc., without any honest attempt at justification or explanation for why the agent should act in such a way. If the agent does so act, she will have acted under some external constraint to her will, imposed by another actor.

⁷ Williams, 2008, pg. 16-18. I added (a) and (b) to two for ease of ease of analysis.

⁸ Bernard Williams, "How Free Does the Will Need to Be?", *Making Sense of Humanity* (Cambridge University Press: Cambridge, 1995.) pg. 4.

⁹ It is not universally accepted that incentive can count as coercion, but I see no reason to assume Williams does not think it can. For a discussion see John W. Chapman and J. Roland Pennock ed. *Coercion* (Chicago: Aldine, 1972.) Chapters 2,3,4 and 13.

The above definition is admittedly incomplete, but it is sufficient to set us up to see how the two extreme theories of reasons for action relate to coercion and the trouble Williams runs into regarding reasons for action and coercion.

One Extreme: The Sub-Humean Model

The most extreme formulation of internal reasons theory, a desire-based model for reasons, holds that an agent has a reason to act only if she has a current desire to perform that action or a desire which performing that action would lead to the fulfillment of. On this model, which Williams refers to as “the sub-Humean model”¹⁰ (because he takes it to be a caricature of Hume’s position), an agent has no grounds upon which to compare reasons beyond surface drives, desires and inclinations. Another agent cannot possibly take the role of an advisor to criticize an agent’s reasons under the sub-Humean model. Even something as seemingly obvious as factual error cannot count in support of undermining an agent’s desire-based reason. Because an advisor cannot criticize an agent’s reasons at all, there is no danger of coercive persuasion using the sub-Humean model. Since the advisor cannot appeal to the agent’s reasons when giving advice, there is no possible way that an advisor *could* coerce an agent, or think she could rightfully coerce an agent into acting on a reason based on the sub-Humean model.

According to Williams, the sub-Humean model represents an overly simplistic model of reasons and overly relativizes reasons to an agent’s whims. Williams rightfully believes that when we say an agent has a reason, we mean more than just that the agent *feels* like doing something. Consider, as Williams does, an agent who wants a glass of gin, but mistakenly believes the glass of petrol in front of her contains gin. An advisor, adhering to the sub-Humean model of reasons, cannot even say of the thirsty agent that she does not have a reason to drink the petrol in front of her.¹¹ The sub-Humean model does not

¹⁰ Williams, 1981 pg. 102.

¹¹ Williams, 1981, pg. 102.

require the agent to reflect at all to have a reason. Williams claims that under the sub-Humean model, an agent can have a reason regardless of factual error, internal inconsistency or similar worry. According to Williams, then, the trouble with the sub-Humean model is that it lacks *any* normative force.

Another Extreme: Williams' Conception of External Reasons.

Whereas internal reasons theory does not just aim to describe internal reasons, but also to exclude external reasons, an external reasons theory seeks a way to add external reasons to a theory of deliberation along with internal reasons. I take Williams' conception of external reasons to represent an extreme version of external reasons, just as Williams meant his description of the sub-Humean model to capture the extreme on the internal reasons side.

According to Williams, an external reasons theory entails that an agent could have nothing in her S to give her a reason, but still be obliged to act on that external reason. Even if she does have something in her S that aligns with that reason, but something about the situation she finds herself in makes it so she does not have an internal reason, she can be held responsible for not acting on that external reason. An agent cannot escape being held responsible for not acting on an external reason when that reason applies. For example, consider, again, the case of Owen Wingrave. If Owen does have an external reason to respect his family, then his father is right to expect him to join the army. The fact that Owen's S contains nothing that would lead him toward joining the army does not factor into whether or not he should.

External reasons, then, carry a particularly strong and accessible normative force in which what an agent has reason to do need not depend on anything about that particular agent. The problem with the normativity of external reasons, as Williams sees it, is that nothing about the agent's perspective or desires need be considered in deciding how the agent should act. Even if she had given her desires deep

and reasonable consideration, making them much more than mere whims, Williams does not think an advisor who appeals to external reasons would have to take those desires into consideration.

The problem, Williams believes, with the strong normativity he attaches to external reasons lies in how that normativity bears on whether or not discourse is coercive. When an advisor finds herself confronted with an agent who refuses to take the advisors' recommendations into consideration, the advisor has no possible persuasive recourse but coercion. However, this assertion of Williams' does not follow from his description of external reasons alone. What an advisor appeals to in persuasion does not necessarily determine how she has to go about performing that persuasion. Consider, again, the example of Owen Wingrave. Although Owen's father believes Owen has a reason no matter what Owen's S contains, Owen's father could still attempt to disclose why joining the army is important. Owen's father could talk about family history, pride and honor. He could recount nostalgic war stories and so on, in an apparently non-coercive attempt to convince Owen to act. I will argue that Williams' conception of external reasons only necessitates coercion if you attach certain attributes to them limiting persuasive options for the advisor.

To defend this claim I have to go into significantly more depth on external reasons than I did on the sub-Humean model of reasons. Williams' claim about the relationship between external reasons and coercion becomes clear, in the light of how Williams' external reasons fit into the rest of his mature philosophical position, specifically his critique of the morality system. It is this connection which makes fully intelligible why he believes external reasons must intertwine with coercion.

The Morality System

Williams displays in his writing great insight into the subtleties of ethical interaction and the realities of the human condition, in particular, realities which he believes traditional morality hides from us or at least encourages us to ignore. He associates morality with "the morality system", a system which he believes is riddled with unhelpful and possibly dangerous myths which we would be better off

without.¹² It's important to note that Williams uses the term "moral" and "ethical" in different ways. He contrasts traditional morality with the more general considerations of ethics.

...morality should be understood as a particular development of the ethical, one that has a special significance in modern Western culture. It particularly emphasizes certain ethical notions rather than others, developing in particular a special notion of obligation, and it has some peculiar presuppositions. In view of these features it is also, I believe, something we should treat with a special scepticism.¹³

The morality system is a subset of ethics which Williams believes captures how we westerners generally view the term "moral." Because this is a paper about Williams, when I say "moral" I will be referring to his usage of the term, and when I say "ethical" I will, as he does, refer to more general structures of how people engage the world. I will not try to give a complete analysis of the morality system. Instead, I will merely focus on the attributes which the morality system shares with external reasons to show how external reasons fit into Williams' morality system narrative. The morality system is most notably defined by a notion of inescapable obligation. The particular brand of obligation which Williams associates with the morality system is colored by a notion of purity, utter voluntariness, a particularly strong notion of blame and a narrow conception of rational persuasion.

Inescapable obligation is the most significant defining feature of the morality system. Obligation is not unique to the morality system. In the everyday notion of obligation, an obligation to do something is merely one consideration among many. Williams considers obligation under the morality system unique because of its inescapability. No matter what an agent's S contains or how little she wants to be a part of it, under the lens of the morality system, no agent can escape being held responsible for acting on moral dictates.¹⁴ While inescapable obligation most significantly shapes the character of the morality system, a number of other attributes color obligation, allowing it to be so strong that moral considerations trump all other considerations.

¹² Bernard Williams, *Ethics and the Limits of Philosophy* (Harvard University Press: Cambridge, 1985), pg. 174.

¹³ Williams, 1985, pg. 6.

¹⁴ Williams, 1985, pg. 178.

Part of why, for Williams, morality can have this special, inescapable, obligatory force is due to the idea of the purity of moral dictates. “The purity of morality [is] its insistence on abstracting the moral consciousness from other kinds of emotional reaction or social influence.”¹⁵ The implication is that moral dictates operate within a type of void and nothing in the agent’s S need keep an agent from acting on moral considerations.¹⁶

Relatedly, adherents to the morality system, as Williams sees it, consider an agent’s actions utterly voluntary. Accordingly, I will call this the “myth of utter voluntariness.” Although related to purity, this myth more specifically holds that an agent can have, through the power of rational thought, total control over her actions. The voluntariness is “total and [cuts] through character and psychological or social determination, and allocate[s] blame and responsibility on the ultimately fair basis of the agent’s own contribution, no more and no less.”¹⁷ Williams sees the morality system as holding that the decision to act or not to act on moral dictates is always entirely in the agent’s power. Not only does the morality system assume an agent has access to moral reasons, it assumes that she has un-fettered access to and control over her own internal motivations. This means that an agent can be expected, under the morality system, to overcome all unconscious drives desires and inclinations as well as circumstantial particulars when making a decision. In reality, any number of considerations such as sub-conscious drives or biological limitations could go into not only determining the action she does take, but the range of possible actions she could take.

Due to the myth of utter voluntariness, morality carries with it a special and dangerous form of “focused blame.” Williams calls blame under the morality system “focused” because it focuses on the

¹⁵ Williams, 1985. Pg. 195.

¹⁶ Kant is Williams’ major target in his critique of the morality system, but this may not be a fair characterization of Kant’s position. See Robert Loudon, “The Critique of the Morality System” in, Alan Thomas ed. *Bernard Williams*. (Cambridge University Press: Cambridge, 2007) pg. 113-126.

¹⁷ Williams, 1985, pg. 194.

particular act without any consideration of the broader context in which the agent acts.¹⁸ Since an agent's actions are seen as entirely in her control-nothing can factor into moral dictates other than specifically moral considerations and the agent cannot escape responsibility for acting on moral reasons-this special unreflective blame is justified under the morality system.¹⁹ Dictates of the morality system are such that no matter what an agent's circumstances, we can hold her responsible for not acting on those dictates.

Finally, and most importantly regarding coercion, Williams believes the morality system entails a misconception of rational persuasion. Under the morality system, due to the myth of utter voluntariness, rational persuasion, narrowly conceived, is the only non-coercive persuasive method, "without [the morality system's] utter voluntariness, there is only force."²⁰ According to this way of thinking, laying out premises and showing that something follows about how the agent should act should convince any rational agent to act. The morality system only allows, then, for two possible methods of persuasion, rational argument, narrowly conceived, or coercion. From the perspective of the morality system, if the advisor does not use strict rational persuasion, she is somehow trying to make the agent act against her will, if not outright coercing her.

While not a complete description of the morality system, a unique obligatory force colored by a notion of purity, utter voluntariness, a particularly strong notion of blame and a narrow conception of rational persuasion capture the crucial nature of how the morality system relates to coercion. These same attributes also help to draw out how Williams' external reasons fit into the morality system.

External Reasons and the Morality System

I will now consider external reasons in greater depth to show how easily they fit into Williams' morality system narrative described above. Once I have delineated the similarities between external

¹⁸ Williams, 1985, pg. 194.

¹⁹ Williams, 1985, pg. 193

²⁰ Williams, 1985, pg. 196.

reasons and the morality system, I can make sense of Williams' claim that external reasons are necessarily coercive.

Williams appears to believe that the only way to conceive of external reasons is as staunchly objective entities which cannot possibly interact with internal reasons in deliberation. The result is a conception of external reasons as extremely strong, accessible and of universally applicable normative force. Although Williams rarely explicitly ties external reasons to the morality system, he clearly attributes to external reasons a special obligatory force and purity. Additionally, according to Williams, it appears that the advisor who appeals to external reasons has to adhere to the myth of utter voluntariness, an unreflective blame and a narrow conception of rational persuasion.

Williams' attribution of a special inescapable obligatory force to external reasons shows up early in his treatment of external reasons in his example of how Owen Wingrave's father interacts with Owen. Williams does not say that Owen's father says the considerations he brings up should factor into Owen's deliberation about how to act. Owen's father makes a claim about how Owen should act no matter what Owen's S may happen to contain. Owen's father does not appeal to anything about how Owen sees the world, his mind is made up before he enters into the discussion, so there is no way Owen can escape responsibility for the reason his father thinks he has.

Williams portrays external reasons as a special class of reasons which cannot even be considered alongside internal reasons in the process of deliberation.²¹ Nothing about an agent's S can go into deciding whether or not or how she should act on an external reason. Williams defines external reasons as reasons to act which do not depend on an agent's perspective or the particular situation she finds herself in. "The whole point of external reason statements is that they can be true independently of an agent's motivations."²² Notice that when Owen's father says he should join the army, he says it knowing full well that nothing in Owen's S would lead him to join the army. This attribute of external reasons' relationship

²¹ Williams, 1981 pg 109.

²² Williams, 1981. pg. 107.

with other considerations bears a striking similarity to the way in which Williams sees moral considerations as pure. External reasons are more than just non-subjective conditions such as the facts of a situation. They are not just special reasons because they depend on non-subjective conditions; they are a special class of non-subjective conditions.

Crucially for the myth of utter voluntariness, external reasons can be universally known by rational agents. Williams sees it as a strength of his internalism that practical deliberation remains an indeterminate process; it is not clear how an agent should act nor what a sound deliberative route is. “Practical reasoning is a heuristic process and an imaginative one, and there are no fixed boundaries on the continuum from rational thought to inspiration.”²³ According to Williams, external reasons cannot share this strength. The way he describes external reasons seems to suggest that an external reason is a clear dictate for action and separated from other considerations in deliberation. If one has an external reason there is no question how or if one should act on it, unless it appears to conflict with another clearly delineated external reason.

Like moral considerations, then, (because external reasons have to be independent of any particular agent’s motivational set, but every rational agent can be expected to come to know external reasons) we can blame an agent for not acting as an external reason dictates. Furthermore, we place blame on the agent without any consideration of any situational particulars or aspects of the agent’s S.

Because of the way Williams characterizes the nature of external reasons, it follows that we can always blame an agent who fails to act on an external reason. No subjective conditions can go into determining whether or how an agent should act on an external reason. Furthermore, given their purity as well as the access agents are meant to have to them, external reasons warrant a special kind of focused blame which gives them a unique coercive potential.

²³ Williams, 1981. pg. 110.
Stephenson

Externalism is indeed external, and the image of blame that can be derived from its account of reasons, in failing to provide any way to engage with an agent's actual motivations, leaves us also without ethical resources. It gives us no way of understanding the difference between a blame that might hope to achieve recognition, and the blame that hopes by mere force to focus on the agent's reasons a judgement that represents in fact only a rejection (perhaps an entirely justified rejection) of what he has done. It leaves us, that is to say, in the condition of moralism.²⁴

This is one of the few places where Williams explicitly acknowledges the connection between external reasons and the morality system. If one blames an agent, one assumes that the agent could have done otherwise and so blame intertwines with the myth of utter voluntariness. When appealing to external reasons, according to Williams, the advisor cannot consider the agent's S or circumstances when deciding whether or not to place blame. Instead, the advisor unreflectively *forces* the agent into adhering to a strict code by placing blame on the agent for failing to act on external reasons.

External reasons, for Williams, act as part of a code the moralist appeals to in the mistaken conception of rational persuasion. One can hope that the power of the better argument will win out because there is a readily accessible code one appeals to which has a pull on all of us regardless of any particular agent's S or circumstantial particulars. Because the moralist assumes that each agent should have access to this code, the initial interaction between the advisor appealing to external reasons and the agent would not necessarily be coercive. The advisor would assume the agent could easily understand the code and proceed to explain it to the agent. However, if the agent cannot be convinced, she may be rationally deficient according to the morality system and, if the reason is important enough, the only way the advisor can get her to act on that reason is by force.

I have elaborated the different attributes of the morality system in order to clarify why Williams thinks moral dictates and external reasons are inescapably obligatory, that is, no matter who the agent is or what circumstances she finds herself in, she cannot escape from responsibility for not acting on applicable external reasons. Because external reasons are obligatory and give clear universally rationally accessible dictates which cannot be adulterated by subjective conditions or circumstance particulars, they

²⁴ Williams, 1995 pg. 44.

trump internal reasons. It is this trumping power which falls out of the tie between the morality system and external reasons and helps to make intelligible Williams' claims about external reasons and coercion.

Making Sense of Coercion in External Reasons

Now that I have shown how Williams melds external reasons into the morality system, I can at least make his claim about external reasons and coercion internally consistent to his own philosophy. Again, coercion for Williams, as I see it, is when one agent tries to get another to act against her own wishes, whether through force, deception, trickery, irresistible incentive etc., without any honest attempt at justification or explanation for why the agent should act in such a way. Consider the scenario where an advisor appeals to external reasons as Williams conceives of them. The advisor cannot appeal to the agent's motivations when attributing an external reason to her. The agent cannot input her motivations in the deliberation, or augment the required action with internal reasons. Certainly it may be the case that the agent's internal reasons align with the external reasons, but this is neither interesting nor dangerous. The scenario that matters is the one in which the agent is not or cannot be motivated to act on the proposed external reason or even see that there is an external reason. Since the reason cannot be augmented to fit the agent's motivations or particular situation and all competent agents are supposed to be able to access and understand external reasons and how to act on them, the stubborn agent must be somehow deficient. The deficiency could be rational, if the agent did not even acknowledge the existence of an external reason. On the other hand, if the agent could see that there was an external reason, but still refused to act on it, she would be deficient in that she was *akratic*.

When appealing to external reasons, the advisor's understanding of the situation is, in an important sense, already determined before seeing the particulars of the situation. If the advisor correctly identifies an external reason, the agent cannot reasonably refuse to act. Rational persuasion, narrowly conceived, the only non-coercive means of persuasion under the morality system, has failed. Since the morality system allows for no form of persuasion other than narrowly conceived rational persuasion, the

advisor has no further recourse to justify why she wants the agent to act. Coercion, then, is both the advisor's only remaining recourse and a perfectly justifiable one. The external reason is still obligatory and, even if the agent suffers from some deficiency which renders her incapable of seeing that reason or how it applies, she still bears responsibility for failing to act on that reason.

Conclusion

The sub-Humean model of reasons and Williams' conception of external reasons represent two extreme ends of a continuum. On one side of the continuum, the sub-Humean model, reasons have no normative force and so an advisor cannot even appeal to them in persuasive discourse. Williams' conception of external reasons captures the opposite end of the continuum. Williams' external reasons have such strong normative force that an advisor need not appeal to anything about an agent's S when dictating action based on external reasons. However, it is only because Williams ties external reasons to the morality system through inescapable obligation, purity, the myth of utter voluntariness, focused blame and a misconception of rational persuasion that external reasons appear highly coercive in persuasive discourse.

Williams tries to find, with his own version of internal reasons, a point on the continuum between the sub-Humean model of reasons and his conception of external reasons. He inserts normativity into internal reasons while trying to maintain that all reasons are internal. This middle ground, according to Williams, allows for criticism of another agent's actions, but also allows for many other avenues of persuasion than his version of external reasons. Because Williams' internalism includes a broad sense of persuasion and a requirement for the advisor to engage the agent's S, he believes it is less likely to fall into coercion than his version of external reasons. Against this, I will proceed in, Chapter II, to explain how easily Williams' middle ground actually could, in fact, fall into coercion. Once I have disclosed the coercive potential in Williams' own internal reasons project, I can show, in Chapter III, a point on the continuum which allows for external reasons that avoid his fears of coercive discourse. I argue that

Williams only attributes coercion to external reasons because he overlooks that point on the continuum which allows for a less severe conception of external reasons than his morality system inspires.

Part II: Two Attempts at a Middle Ground

Chapter II: Williams' Internalism

Williams aims much of his internalism project at making internal reasons as normative as possible while shying clear of his extreme view of external reasons. External reasons, he believes, lead too easily to coercion, but the sub-Humean model does not allow for any criticism of an agent's reasons at all. So, Williams attempts to find a middle ground with a non-coercive conception of reasons which allows for a fairly strong conception of normativity but avoids external reasons. Unfortunately, in the course of navigating this space between extremes, Williams exaggerates the protection normative internal reasons offer against coercion.

Before drawing out the main problem with Williams' conception of normative internal reasons, I must parse out two seemingly different versions of normative internal reasons which appear in his work. I will refer to the two forms as *weak normative internal reasons* and *strong normative internal reasons*. These two versions of normative internal reasons may not be compatible, and it does appear as though Williams prefers the strong version, but I need not reconcile these two threads in Williams' work. For my purposes here, it is enough to point out that neither of his versions of normative internal reasons works as a middle ground. The weak version of normative internal reasons lacks the normativity he set out to incorporate into internal reasons, and strong normative internal reasons do not offer significant protection against coercion.

Throughout Williams' account of internal reasons he makes an agent's having a reason to act contingent on whether or not there is a sound deliberative route from the agent's existing motivations to

the motivation to perform that action. The strong and the weak versions of normative internal reasons both share a notion of a sound deliberative route where an agent may be wrong about her own reasons, to some extent. By opening the possibility that an agent may be mistaken about her own reasons, Williams opens a space for another agent to take the role of an advisor to criticize an agent's proposed reasons. The role of the advisor is crucial for a conversation about coercion because coercion requires at least two actors. An advisor may recognize the agent's motivations and offer advice based on that recognition. Then an advisor can criticize an agent's proposed internal reasons by pointing out aspects of the world, pointing out other motivations the advisor thinks the agent may have or even pointing out the agent's failure to properly imagine the outcome of her actions. All of these considerations go into the sound deliberative route which, like the S and what an agent has reason to do, is extremely, and Williams argues importantly, indeterminate. For an advisor to claim, with any semblance of authority, that she appeals to an agent's reasons when persuading her to act, she at least must make an honest attempt to engage the agent's S.

Both of Williams' versions of normative internal reasons share the above requirements for a sound deliberative route. Neither version, however, actually acts as a satisfactory middle ground between the sub-Humean model of reasons and external reasons when dealing with coercion.

Weak Normative Internal Reasons

In Williams' weak normative internal reasons, the advisor has very little power to criticize the agent's reasons. This conception of internal reasons only comes later in Williams' work, but because he intends it as one of his final clarifying remarks on his internal reasons theory, I believe it warrants some serious consideration.

Williams' weak normative internal reasons appear in "Values, Reasons and the Theory of Persuasion" where he wants to put some constraints on what counts as a sound deliberative route. I have already brought up these requirements in my discussion of "good deliberative assistance", but they also warrant consideration here. According to this view, not just any desire which an agent may happen to have counts as a reason, but if an advisor, adhering to the following requirements, cannot convince the agent she has a reason, the agent cannot be said to have an existing internal reason.²⁵

- (i) The assistant will be truthful, in the sense both of telling the truth and of helping the agent to discover the truth. An application of this is that the assistant will be truthful about his own procedures and motives, with the result that these can be transparent to the agent; the assistant has no hidden agenda in his dealings with the agent.
- (ii) [a] The assistant will try to make the best sense of the agent's S, and, in particular, [b] if there is a conflict between the assistant's and the agent's interpretations of the agent's S, the assistant will have some suitable explanation of the agent's misinterpretation.

Under these requirements, it seems that if the assistant makes an honest effort to convince the agent and the agent remains unconvinced, then she does not have a reason. Although the focus of the above requirements is on the advisor, because these requirements make an agent's having a reason contingent on whether or not an advisor can convince the agent to act on a reason, they say something about an agent's relationship to her own reasons. The agent need not consider the advisor's interpretation, or thoroughly consider her own opinion to genuinely have a reason. Whether or not an agent has an internal reason, based on the above requirements, is thoroughly contingent on how self aware the agent is and how responsive she is to another's criticism.

These weak normative internal reasons back a theory of deliberation which makes it difficult for an advisor to use coercion on an agent when appealing to internal reasons. Because the advisor essentially has to give up on a reason if an agent is difficult to convince, the advisor has no incentive to coerce an agent to act on an existing internal reason. If the advisor has to resort to coercion to get an agent to act, the agent automatically does not have a reason.

²⁵ Williams, 2008, pg. 16-18.
Stephenson

The problem with Williams' weak normative internal reasons is that they bear little practical difference to the sub-Humean model of reasons for action which Williams wants to avoid. Under Williams' weak normative internalism, it appears that an agent has a reason so long as she either has a manifest desire to perform a certain action or can be easily convinced to perform that action. Williams' gin and petrol example can still work to illustrate weak normative internalism. An agent could be easily convinced not to drink the petrol she mistakes for gin simply by the advisor pointing out that the glass in front of her contains petrol. However, if petrol turns out to be the agent's favorite beverage, or she is wildly stubborn and does not want to listen to anything the advisor says, then she still has a reason to drink the petrol in front of her. The only significant difference between this weaker version of normative internal reasons and the sub-Humean model is that an agent may be mistaken about her immediate desires, allowing an advisor this ground upon which to criticize the agent's actions.

An advisor cannot attribute a weak normative internal reason to an agent if persuasion based on the requirements of good deliberative assistance fails. So something as mundane as an agent not wanting to listen to the advisor seems to be enough to discount the advisor's deliberative assistance. The result is that weak normative internal reasons have very little normative force. Williams intends his remarks around the requirements for good deliberative assistance to clarify how agents interact, how they might appeal to internal reasons and even to show how different internal reasons and external reasons are in relation to coercion. Unfortunately, the result of Williams' attempt at clarification is to weaken his normative internal reasons to the point where they are practically indistinguishable from the sub-Humean model of reasons which he set out to avoid.

Strong Normative Internal Reasons

Williams more frequently presents his strong version of normative internal reasons, which also fits better with the arc of his overall account of reasons for action. According to the strong view, an agent

has a reason if a sound deliberative route *exists* from her current motivations to performing that action.²⁶ For Williams, the agent does not actually have to be able to see that route in order for something to be a strong normative internal reason.

Not every desire counts as a strong normative internal reason to act. First, sound deliberation requires that the agent properly exercise her imagination such that she understands the reality of the outcome of the proposed action.²⁷ Sound deliberation, under Williams' strong normative internal reasons, also requires factual accuracy. If an agent wants a cup of gin and mistakenly believes the cup of petrol in front of her is gin, she does not have a reason to drink what's in front of her. Sound deliberation, then, requires that the agent properly understands the facts of her situation. In the gin and petrol case, if the agent paused to take better account of the situation, she would find that she had no reason to drink what was in front of her. So there is not actually a route from her existing motivations, which, according to Williams, always include some drive for factual accuracy, to drinking the gasoline she mistakes for gin.²⁸

I doubt accurately discerning the facts of a situation would always be as simple as noting that a glass contains petrol instead of gin, but by factual accuracy Williams does seem to mean fairly mundane facts about the material world. He gives no requirement, for instance, for an agent to have proper understanding of other agents' perspectives.²⁹

Prudence also comes to play a role in Williams' notion of a sound deliberative route. By prudence here I just mean an agent's knowledge of his own good and his ability to realize it. Seeing and fulfilling basic needs is a straightforward example of prudence. An agent acting with no care towards his future self is an obvious case of imprudence. Although Williams initially denies a requirement for prudence in sound

²⁶ Williams, 2001. pg. 91.

²⁷ Williams, 1981, pg. 105.

²⁸ Williams, 1981, pg. 102. Williams, 2001 pg. 92, Williams, 2008 pg. 111.

²⁹ Williams, 1995. pg. 41-42. As evidenced by Williams' notion of proleptic blame. He says if an agent finds out after the fact that another agent the first agent respects disapproves of the agent's action, then perhaps the actor can be blamed. The actor did not, at the time of the action, have a reason not to act, in spite of being mistaken about how the person he respected would judge the action.

deliberation,³⁰ he eventually admits that in order for another to make sense of an agent's actions, the agent needs some minimum level of prudence.³¹ If the agent's actions seem entirely against the agent's needs, such as starving himself, or attempting suicide, the reasons may be unintelligible to an outsider and not part of a sound deliberative route. It's not at all clear what level of prudence is necessary on Williams' account, but once he incorporates prudential reasons into internal reasons, moral reasons are the only viable candidates for external reasons.³²

Under Williams' strong normative internal reasons, having a reason does not depend on imagination, factual reflection and prudence alone. These three requirements for a sound deliberative route do offer an advisor plenty of ways to criticize an agent's actions, giving internal reasons some normativity. The particular weight of strong normative internal reasons, however, comes from Williams' requirement for *self* reflection. An agent's S is not necessarily transparent to the agent, or at least not immediately so.

Supporting the complexity Williams attributes to the agent's relationship to her own strong normative internal reasons is Williams' treatment of the role of the subconscious on reasons for action. Although his treatment of the subconscious is quick and unclear, I can gain from it some insight into an agent's relationship with her own reasons. Williams claims that in order for an agent's subconscious motivations to have a bearing on her reasons, the subconscious motivations need a direct relationship with the action in question.³³ I take Williams here to mean that the motivation, if manifest, would have to have a major influence on what an agent has most reason to do. This does open some room to criticize the wildly un-self-reflective for not adequately knowing their own motivations, but it does not represent the full extent to which an agent can be mistaken about her own reasons.

³⁰ Williams, 1981. pg 105-6.

³¹ Williams, 2008, pg 111.

³² Williams, 1995, pg. 36.

³³ Williams, 1981. pg. 103.

In response to a paper by John McDowell on Williams' first argument against external reasons,³⁴ Williams makes it clear that the agent does not even have to be able to take the sound deliberative route in order to have a reason.

But this does not imply, as McDowell's objection suggests, that the agent should be able to conduct the relevant deliberation in fact. Perhaps some unconscious obstacle, for instance, would have to be removed before he could arrive at the motivation to ϕ . Someone who claimed something of this sort would have to make good on the claim; he would have to be able to make some distinction between, for instance, removing blocks to the expression of A's existing S, and adding to A's S.³⁵

Here Williams clearly states that an agent's S is not entirely transparent to the agent, but just because an agent may not be able to see she has a reason does not alone mean she does not have one. A gap yawns between what the agent thinks she has reason to do and what she does, in fact, have reason to do. This gap helps to illuminate how internal reasons can have strong normativity.

Williams solidifies this position as he continues to develop his theory of internal reasons in "Values Reasons and the Theory of Persuasion"

[The agent] may do everything that could conceivably be demanded of a rational agent, but fail through bad luck, in particular through non-culpable ignorance, and in that sense he will not have done what he actually has most reason to do (this is the sense in which 'a sound deliberative route' if it is to provide a sufficient condition of what the agent has most reason to do, must consist of more than simply the agent's deliberating as best he can in the circumstances.)³⁶

Even if the agent tries to take the sound deliberative route, doing everything in his ability to figure it out, he could fail to act on his reasons. The result of all of this is that the sound deliberative route is extraordinarily indeterminate, making what an agent has reason do also indeterminate. This allows for strong normativity because it may not be clear to an agent in any given situation what she has reason to do. Under weak normative internal reasons, on the other hand, so long as the agent deliberates as well as

³⁴ John McDowell. "Might there Be External Reasons?", in J.E.J. Altham and Ross Harrison ed. *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. (Cambridge University Press: Cambridge, 1995)

³⁵ Williams, 1995, pg. 188

³⁶ Williams, 2008, pg. 113.

he can, the advisor has to concede that the agent correctly ascertains his own reasons. In this way, an agent can more easily access her own reasons under the weak normative internal reasons interpretation.

To illustrate how an agent might be wrong about her internal reasons when they exhibit strong normativity, consider an example which T.M. Scanlon uses to illustrate this point.³⁷ Paraphrasing Scanlon's example, an agent, O'Brien, absolutely loves proper etiquette, but due to some deficiency cannot see how rude he is in most situations. O'Brien's S is thoroughly wrapped up with being polite. He is so open about his love of etiquette that it would require little effort on the part of the advisor to see that O'Brien's S points to reasons to be polite. No matter how hard the advisor tries, however, she cannot get him to change his ways because he remains convinced he is being polite. The advisor adheres to all of the requirements of good deliberative assistance, maybe even getting guests from O'Brien's most recently botched dinner party to say how offended they were, but O'Brien just does not understand. O'Brien is quite obviously mistaken about what he has reason to do. The actions he performs are contrary to his S, but he somehow misses the disconnect between what he wants and how he acts. It is in this sense that it can seem to an advisor that an agent has an internal reason but remains incapable of seeing it.

I have illustrated what Williams requires of an agent to perform sound deliberation and shown how the requirements of sound deliberation open space for an agent to be wrong about her own reasons. This space allows an advisor much more critical power when appealing to strong normative internal reasons than weak ones. In spite of this weighty normativity, Williams does think his strong normative internal reasons carry some protection from coercion, although not as much as weak normative internal reasons. I interpret his position on this point as follows.

Many opportunities exist for coercion when trying to change something in an agent's S to give her a new reason to act. To home in on how Williams might think strong normative internal reasons

³⁷ T.M. Scanlon, *What We Owe to Each other* (The Belknap Press of Harvard University Press: Cambridge, 1998) pg. 366.

maintain certain protections against coercion, I focus on when an advisor actually tries to appeal to an existing internal reason. To credibly claim to know the agent has an internal reason, the advisor has to have engaged the agent to glean some intimate knowledge of the agent's S.

In order to clarify how Williams thinks strong normative internal reasons carry some protection against coercion, consider, again, his requirements for good deliberative assistance.

- (i) The assistant [or advisor] will be truthful, in the sense both of telling the truth and of helping the agent to discover the truth. An application of this is that the assistant will be truthful about his own procedures and motives, with the result that these can be transparent to the agent; the assistant has no hidden agenda in his dealings with the agent.
- (ii) [a] The assistant will try to make the best sense of the agent's S, and, in particular, [b] if there is a conflict between the assistant's and the agent's interpretations of the agent's S, the assistant will have some suitable explanation of the agent's misinterpretation.

Under weak normative internal reasons, if an agent remains unconvinced after good deliberative assistance, she does not have a reason. Under strong normative internal reasons, an advisor merely has to perform good deliberative assistance in order to credibly claim she appeals to an internal reason.

Above I claimed that Williams considers coercion to be when one agent tries to get another to act against her own wishes, whether through force, deception, trickery, irresistible incentive etc., without any honest attempt at justification or explanation for why the agent should act in such a way. If you apply the requirements for good deliberative assistance to strong normative internal reasons, the following picture falls out.

Regarding (i), the advisor may have difficulty being entirely deceptive about her own motives because she has to figure something out about the agent's S in order to credibly claim that she appeals to the agent's reasons. (ii.a) then, is already taken care of, so long as the advisor wants to appeal to the agent's reasons because knowing an agent's reasons requires that the advisor genuinely tries to understand the agent's S.

If a disagreement arises in persuasion, the advisor cannot simply say “well because you have to”, she will most likely have a suitable explanation for her position because in most cases she should assume the agent has authority over her own S. So, (ii.b), will likely be fulfilled as well. Williams thinks, and rightfully so, that if there is a disagreement about the agent’s reasons, the burden of proof rests on the advisor.³⁸ Even if the agent remains unconvinced and the advisor genuinely wants the agent to act, the advisor still has many possible persuasive methods to explore.

Aside from coercion, if the advisor is convinced that the agent is simply wildly un-self-reflective and she knows the agent intimately, she may attempt to change the agent’s mind through conversion or emotional appeal and still be appealing to the agent’s internal reasons. Williams does not want to exclude forms of persuasion other than traditional rational persuasion from the interaction between the advisor and the agent; in fact, he does not think a hard distinction between rationality and persuasion has any place in practical reasoning.³⁹ Williams believes the interaction where an advisor helps an agent to discover an existing internal reason can include emotional appeals and other more creative means than rational discourse, narrowly conceived, without falling into coercion.

The weak form of normative internal reasons Williams presents is unsatisfactory as a middle ground due to its resemblance to the sub-Humean model which Williams wants to avoid. In his strong version of normative internal reasons, Williams wants to maintain the normativity which comes with external reasons, but keep reasons relative to an agent’s S, maintaining all reasons as internal reasons. He sets a better middle path with his strong normative internal reasons between the sub-Humean model for reasons and the staunch externalism he maligns. Strong normative internal reasons support a theory of reasons which allows for one agent to advise and critique another, but still requires each agent to engage the other’s S and disclose facts about his own S. This necessary engagement with and disclosure of S’s, according to Williams, results in some protection against coercion when compared to external reasons.

³⁸ Williams, 2008, pg 115-116.

³⁹ Williams, 2008, pg.115.

Unfortunately, Williams' strong normative internal reasons much more closely resemble a more plausible form of external reasons than Williams lets on and do not result in significantly more protection from coercion.

Coercion with Strong Normative Internal Reasons

Because an agent's S lacks self-transparency under Williams' strong normative internal reasons, the possibility always remains open that she could simply be wrong about her own reasons. While an advisor's knowledge of an agent's S will always remain imperfect, space remains open for the advisor to be correct and the agent to be incorrect about the agent's reasons. This space makes coercion a live possibility with Williams' strong normative internal reasons.

Let me now examine a very specific scenario in order to compare the coercive possibilities in appeals to external reasons and strong normative internal reasons. In this scenario, the advisor and the agent have to disagree about how the agent ought to act. The advisor thinks the agent should do something the agent does not think she has good reason to do. In order to know an internal reason, the advisor has to engage the agent's S. So when trying to persuade an agent to act on that reason, if the advisor is to have any credibility in claiming she wants the agent to act on an internal reason, she has to know the agent intimately or discuss the situation with the agent. To draw the analogue between this scenario and the scenario where the advisor has to coerce an agent to act on an external reason, the disagreement has to have come to an impasse, i.e. the advisor has engaged the agent's S, tried to convince the agent she has an internal reason to act, but the agent remains unconvinced. The agent does not see any reason to act and the advisor remains convinced, based on what she makes of the agent's S, that the agent ought to act. Remember, I am talking about appeals to existing internal reasons because obvious avenues exist for coercion when trying to *change* an agent's S.

Three scenarios could come to pass in a disagreement like the one above where the interlocutors have reached an impasse.

- 1) The advisor could throw up her arms and give up trying to convince the agent.
- 2) The advisor could indefinitely continue trying to show the agent her error.
- 3) The advisor could try other means to get the agent to act, such as coercion.

I can discard (1). To draw the parallel between coercion in appeals to internal and coercion in appeals to external reasons I need to assume that the advisor has some strong commitment to getting the agent to act. If it did not matter to the advisor if the stubborn agent acted or not, even if she thought the agent had an external reason to act, the advisor would not force the matter. (2) appears absurd. At some point the disagreement will have to stop and either move to (1) or (3), so I have to focus on (3). The question, then, is whether or not Williams' strong normative internal reasons hold any special attributes that make coercion exceptionally unlikely in (3). I use the remaining space in this section to show that they do not.

Williams admits in his comments on reasons for action that an advisor can have a place in an agent's deliberation. The fact that Williams wants to make reasons normative commits him to allowing that an advisor can criticize an agent's proposed reasons and advise her toward reaching an internal reason which she may not have been able to reach alone. The reasons are internal, based on the agent's S and a sound deliberative route. In order to genuinely enter into an agent's deliberative process, the advisor must, as Williams says repeatedly, consider the world from the agent's perspective.⁴⁰ The advisor must ask herself "If I were you, what should I do?"

As I said above, an agent's own reasons are not necessarily clear to her and even if she cannot see she has a reason, the advisor does not have to concede that the agent does not have a reason. I think it is safe to assume that in many cases the advisor should defer to the agent about her own reasons. On the other hand, since an agent's S is not necessarily clear to her, it can, at times, be easier to see what makes sense for another agent than for herself. The scenario of an advisor trying to persuade a recalcitrant agent,

⁴⁰ Williams, 2001, pg. 92.
Stephenson

then, is interesting because two agents with fuzzy knowledge of reasons face off. Neither can be sure that the other is wrong, but a spark of doubt could always remain on both sides. In this way, no matter how the agent protests, if the advisor knows her well enough, she may insist that the agent still has an internal reason to act. The space for disagreement, then, is infinite. The advisor and agent disagree about how the agent should act, they have both made their points, disclosing perspectives, but they both remain recalcitrant. If the action matters to the advisor, why would she be more likely to give up than to search for other means of persuasion? And, more to the point, does the fact that these reasons are internal protect against coercion as one of those other means? My answer, of course, is no.

The agent does not have to be able to take the sound deliberative route, so the fact that she remains recalcitrant does not mean she does not have a reason. Williams presents coercion as when one agent tries to get another to act against her own wishes, whether through force, deception, trickery, irresistible incentive etc., without any honest attempt at justification or explanation for why the agent should act in such a way.

For ease of understanding, consider, again, the requirements for good deliberative assistance which appear to offer some protection against coercion with strong normative internal reasons.

- (i) The assistant [or advisor] will be truthful, in the sense both of telling the truth and of helping the agent to discover the truth. An application of this is that the assistant will be truthful about his own procedures and motives, with the result that these can be transparent to the agent; the assistant has no hidden agenda in his dealings with the agent.
- (ii) [a] The assistant will try to make the best sense of the agent's S, and, in particular, [b] if there is a conflict between the assistant's and the agent's interpretations of the agent's S, the assistant will have some suitable explanation of the agent's misinterpretation.

If an advisor follows (i) and (ii.a) but the agent remains unconvinced and the action is important enough to the advisor, she may give up on (ii.b). There is infinite space for disagreement so, at some point, since the disagreement must end, the advisor may coerce the agent and actually be coercing the agent to act on an internal reason. I say it is coercion because even though the advisor went through the process to avoid coercion, eventually, if she actually gets the agent to act, she would have to step outside of these bounds

of coercive free discourse. She may stop trying to justify her persuasion to the agent, and nothing about the nature of strong normative internal reasons suggests that she cannot.

Conclusion

Williams finds unsatisfactory both the sub-Humean view of reasons and the extreme view of external reasons he maligns. He tries, with his internal reasons theory, to strike a middle ground between the two extreme views of reasons in an attempt to find a way to criticize an agent's reasons while avoiding the dangers he sees in external reasons. Two separate versions of normative internal reasons fall out of Williams' attempt at a middle ground: strong and weak normative internal reasons. The weak version too closely aligns with the sub-Humean model, so it fails as a middle ground because it lacks normative force. The advisor barely has space to advise or criticize an agent's reasons, which is a scenario Williams wants to avoid with internal reasons.

Williams, with his strong normative internal reasons, leaves considerable space open for the advisor to coerce the agent to act on an existing internal reason. Since sound deliberation requires factual accuracy, consistency, some level of prudence and a fairly deep level of self reflection, if the agent does not appear to live up to one of these requirements, the advisor has many avenues to simply browbeat her into acting on a perceived internal reason. The strong version of normative internal reasons, then, does not offer any *special* protection against coercion.

By appealing to strong normative internal reasons, an advisor does not actually appear to be especially unlikely to use coercion, but strong normative internal reasons may still offer some protection against coercion, as compared to Williams' treatment of external reasons. Because the advisor has to engage the agent to credibly claim she is trying to get the agent to act on a strong normative internal reason, she will at least have to begin to discuss the situation with the agent, even if that discussion only has to go so far. My task, then, is to now disclose how Williams' conception of external reasons as part of the morality system and necessarily coercive is misguided. Nothing about the nature of external reasons

Stephenson

suggests that an advisor must use coercive persuasion when appealing to external reasons. Once we have in view a reasonable account of external reasons, Williams appears to have to either accept the model of weak normative internal reasons which so closely resembles the sub-Humean model, or concede that his strong normative internal reasons are no less likely than external reasons to lead to coercion.

Chapter III: A Reasonable Conception of External Reasons

The Primary Content of External Reasons

Williams famously rails against the “fictions” of the morality system. As he develops his critique of external reasons, he appears more and more to view external reasons as just another fiction, a character in the larger story of the morality system, entailing the same dangers as other moral considerations. While the attributes which Williams’ external reasons share with the morality system could be part of an extreme externalism, nothing about the nature of external reasons suggests that these attributes are necessary. By pinning down exactly what makes a reason external and showing how little that constrains how we can conceive of external reasons, I can show why Williams’ claim that there is a strong relationship between external reasons and coercion is unfounded.

I will call what makes a reason internal or external that reason’s *primary content*. A reason’s primary content is not necessarily what makes it a reason and certainly not what dictates how the agent should act. In the gin and petrol case, the desire for gin, assuming adequate self understanding on the agent’s part, makes up the primary content of the reason. Whether or not it actually is a reason and how it should manifest in action depends, among other things, on non-subjective conditions, in this case facts. The facts of the situation coupled with the agent’s S make the desire for drinking what is in front of him not a reason. On the internalist account, if the thirsty agent also loves drinking gasoline, he very well may have a reason to drink from the cup he thought was full of gin, he would just be pleasantly surprised by

gasoline, his favorite beverage.⁴¹ Giving content to external reasons, as Williams is fond of pointing out, is not so easy, but I will have to try in order to make intelligible the alternate picture of external reasons I wish to disclose.

For an external reason, what makes it a reason cannot depend on any agent's S; that is the crucial line that we cannot cross if we want to continue talking about external reasons. Williams correctly believes that the *primary content* of an external reason has to be pure.⁴² He is, however, mistaken about how that content has to factor into deliberation. Just as Williams maintains reasons internality while giving them a palatable level of normativity, I can maintain some reasons externality while requiring *some engagement* with the agent's S. Williams mistakenly believes that just because an external reason cannot depend on an agent's S, the agent's S cannot be considered alongside external reasons in deliberation. External reasons, in fact, merely require that the reasons themselves do not depend on subjective conditions for their bindingness. The purity of content says nothing about how external reasons have to factor into an agent's deliberations. So regarding their content, external reasons are necessarily pure, but Williams' external reasons seem to require purity and primacy in deliberation as well.

Williams never argues that external reasons have to trump other considerations, he merely assumes a picture of deliberation in which they do. The trumping power of external reasons falls entirely out of Williams' portrayal of them as part of the morality system. For Williams' claim that external reasons have a special relationship with coercion to be correct, a reason's externality would have to require that external reasons also have the other attributes which tie them to the morality system: special obligatory force, a more complete sense of purity, a requirement for adherence to the myth of utter voluntariness, a unique focused blame and a limited conception of rational persuasion.

⁴¹ As I mention above, Williams does eventually concede that some level of prudence is necessary for a sound deliberative route. Just what level of prudence remains unclear, so, depending on how one makes intelligible a notion of internalism which includes prudential considerations, this example may or may not hold in the final analysis.

⁴² Williams, 1981, 107.

The pure primary content of external reasons does not say anything about what type of persuasive method an advisor has to use to convince an agent that an external reason applies to her. Just because the entity is pure does not mean it has to be part of a code easily accessible to all rational agents. Because purity requires neither unique inescapable obligation nor a conception of the agent as capable of utterly voluntary action, the particular focused blame of the morality system is simply not *necessarily* connected with external reasons. My point here is merely that the only necessary attribute Williams seems to touch on for external reasons is that they have to be pure, and as I said, they only have to be pure in their primary content which says nothing of their application.

It is not enough to simply point out that Williams' conclusions about the coercive applications of external reasons do not follow from their purity. I must make intelligible a picture of external reasons as merely one part of the deliberative process; a picture in which external reasons can be considered alongside internal reasons in deliberation and do not require all of the other attributes which tie external reasons to the morality system. Once I have set out this alternative model, since the coercive properties Williams attaches to external reasons do not follow from those reasons' property of externality, I will have succeeded in showing that external reasons do not bear a necessary connection with coercion.

Escapability and External Reasons

It's difficult to imagine what I mean by an external reason merely factoring into deliberation alongside internal reasons and other considerations because I think we often assume, as Williams does, that some sort of special inescapability has to be attached to objective dictates for action. I propose we have to think about objectivity, particularly as it relates to reasons for action, entirely differently than that. What an agent ultimately should do will most likely depend on a combination of many reasons, perhaps internal, external or both. I envision external reasons merely factoring into deliberation in a similar way as

they do in the naturalistic account developed by Philippa Foot.⁴³ Whereas Williams' account of external reasons forces reasons into a strict easily accessible code, the account I have in mind allows that external reasons may be difficult to clearly understand and merely act to guide our action rather than act as standalone dictates to action.

First, I need to clarify what I mean by naturalism. I do not mean that external reasons manifest out of nature and can be ascertained in the same way as natural scientific facts. The sort of naturalism I want to point to as an alternative way to conceive of external reasons has been expertly discussed by Philippa Foot. Late in Foot's career she tries to give content to external reasons, but avoids the traps Williams thinks external reasons entail. The main thrust of Foot's argument in *Natural Goodness* is that we should re-imagine practical reason in such a way that it can include objective ethical considerations alongside other considerations, such as desire based reasons.⁴⁴ The result of her analysis is to include a type of external reason within a complex deliberative process, instead of requiring external reasons to stand alone in deliberation.

I can hold Foot's position up as an alternative to Williams'. Foot's position will help to disclose an intelligible conception of reasons, maintaining their primary content as external, but still requiring the advisor engage the agent's S when determining a practical should. Outlining Foot's alternative will act to defeat Williams' claims about the necessary connection between external reasons and coercion.

Couching external reasons in Foot's naturalism makes it much easier to see how the scope of possibilities for external reasons could be broader than Williams allows. External reasons, conceived on Foot's grounds, are merely guides which help shape a good human character. How one should actually act in any given situation depends on many different considerations.

⁴³Philippa Foot, *Natural Goodness*, (Oxford University Press: Oxford, 2001).

⁴⁴ Foot, pg. 13.

For Foot, “moral defect is a form of natural defect not as different as is generally supposed from defect in sub-rational living things.”⁴⁵ Moral defect can be ascertained, according to Foot, based on what she calls, building on Elizabeth Anscombe’s terminology, “Aristotelian necessity.”

We invoke the same idea when we say that it is necessary for plants to have water, for birds to build nests, for wolves to hunt in packs, and for lionesses to teach their cubs to kill. These ‘Aristotelian necessities’ depend on what the particular species of plants and animals need, on their natural habitat, and the ways of making out that are in their repertoire. These things together determine what it is for members of a particular species to be as they should be. And for all the enormous differences between the life of humans and that of plants or animals, we can see that human defects and excellences are similarly related to what human beings are and what they do.⁴⁶

The idea of an Aristotelian necessity invokes a shared ground among human beings. There are some things that simply by being human, an agent has reason to do. Aristotelian necessities are not clear dictates for action, but rather general guidelines for what it is for humans to do well. Certainly, given our reasoning capabilities, knowing what it is for humans to do well is much more complex than noting that a lioness should teach her cubs to hunt or an oak should have strong roots. For instance, promising is the type of thing that human beings just cannot get along without. Saying “One should keep promises” does not give a clear dictate for action under this conception. What it does, instead, is tell us that a person who never keeps promises is, in some sense, defective. That person is defective just as we would say that an oak which bends and sways in the wind is defective.⁴⁷

To better understand the possible richness of how Foot’s external reasons factor into deliberation, consider another good candidate for an external reason based on Aristotelian necessity. We can say that people have a reason to help take care of their aging parents, but this does not give an advisor a justification for chastising an agent for not helping her ailing father no matter what the context or what the agent’s S might contain. If the agent’s father raped her mother and never had any contact with the family, our intuition is that although humans have reason to honor their parents, in this scenario the advisor would be wrong to think that the fact that the ailing rapist is the agent’s father should factor into the

⁴⁵ Foot, pg. 27

⁴⁶ Foot, pg. 15.

⁴⁷ Foot, pg. 46

deliberative process in the customary way. How that reason factors into deliberation, shared or individual, depends on something specific about that agent's context and history.

Foot's picture of external reasons allows for a much more complex deliberative process than Williams' picture, but still allows for the externality of some reasons. Sometimes, as in promising, external reasons may factor into deliberation. Promising could be a factor in deliberation if, for instance, you promised to help a friend move but on the way to help your friend you had a chance to further some cause which you held dear. The fact that you made a promise should factor into your deliberation, but so might the importance of the cause or how fleeting this specific chance to help, as well as how much the friend needs you and so on. The external reason that people should keep promises would simply weigh in your deliberation, but not straightforwardly tell you what you should do or advise another to do. At other times, as in the rape case, the external reason may be entirely silenced by other considerations. Foot's point is that ethical reasons, which are in this case external, should factor into practical reason, not be thought of as separate from practical reason generally or a separate process of practical reasoning. Just when an external reason should apply and exactly what force it should have is a worthwhile question without, contrary to Williams' version of external reasons, a clear or easily understood answer. An external reason clearly bears no special weight due to its externality alone.

Foot's reasons based on Aristotelian necessity maintain a primarily pure content. The reasons are not based on any agent's subjective motivational set, but they fit into deliberation in such a way that they do not necessarily trump the agent's S. Because it is not clear whether an agent ought to act on any external reason, the advisor cannot justifiably force that reason on the agent. Foot's position represents one possible alternative way to couch external reasons without conceding the strong connection Williams thinks external reasons have with coercion.

Nothing about a reason's externality suggest that it has to trump other considerations. One should not just assume an external reason's obligatory force. Foot's account represents an alternative way for

external reasons to factor into deliberation which maintains their purity in the necessary sense I described above. Merely allowing that internal reasons could factor into deliberation alongside external reasons breaks Williams' necessary connection between external reasons and coercion.

By disclosing an alternative picture of how external reasons can factor into deliberation I have shown that there is no necessary connection between appealing to external reasons and coercing an agent to act. If an agent refuses to act on an external reason, the advisor may rightfully ask "Why?" Given that the agent responds with good reasons, the advisor ought to concede that perhaps she is mistaken about how that external reason applies in this particular circumstance. Merely because an advisor knows an external reason does not mean she knows what an agent practically should do in any given situation without knowing the particular applicable aspects of the context and the agent's S. The agent and the advisor could still endlessly disagree about external reasons, but that leaves the situation in a similar position as Williams leaves normative internal reasons. In spite of having already broken Williams' necessary connection between external reasons and coercion, I would like to take one more step to fully disclose how narrowly Williams conceives of external reasons by showing that even the epistemological model we use to talk about knowing external reasons could offer some protection from coercion.

Knowing External Reasons

When we start talking about external reasons as natural universals in humanity, there may be a temptation to reduce external reasons to natural law. Although an external reason being akin to a natural law still says nothing about a reason's obligatory force, it may incline an advisor towards browbeating in shared deliberation. Appeals to external reasons may still seem more likely to be coercive than appeals to internal reasons because an advisor can be sure of exactly what external reasons there are and how they should factor into deliberation. I believe this reductionism is severely mistaken. Given the complexity of figuring out Aristotelian necessities, this reductionism is an unlikely epistemological candidate for how

Stephenson

one comes to know an external reason. Before I can fully sever external reasons from the morality system and remove all doubt that there is a necessary connection between external reasons and coercion, I have to briefly sketch an alternative epistemological model for knowing reasons.

How we come to know reasons does not necessarily dictate how we appeal to those reasons in persuasion. However, certain epistemologies may offer some general protection against coercion. Rather than reducing external reasons to natural law, a model which fits well with the clear dictates for action required by the morality system, I would like to propose that a fallibilistic/perfectionist model would fit much better with the model of external reasons I have set out.

Fallibilism does not suggest there is no objective knowledge, in this case of reasons. Instead fallibilism holds that we can never be sure we have obtained absolute knowledge of reasons. Put simply, we could always be wrong. Although not the same, I think fallibilism gestures at a perfectionism. Not only is the fallibilist open to rebuke because there is no easy way to gain knowledge of reasons, but as a perfectionist she never assumes she has gained that perfectly complete knowledge. She always seeks to enrich what she thinks she knows of reasons, possibly even seeking rebuke to develop and strengthen her understanding. Given this epistemological model, it would seem that the fallibilist/perfectionist externalist may be more likely to reexamine her own assumptions rather than to coerce an agent to act on them.

Some of my claims may yet seem tenuous, but I want to suggest how far one can push external reasons toward internal reasons while maintaining those reasons externality. Based on the above formulation, not only have I severed Williams' necessary connection between external reasons and coercion, but, without giving up on external reasons primary content, I have indicated a conception of external reasons according to which an advisor's knowing and applying external reasons requires engaging the agent's S.

Under a fallibilistic/perfectionist epistemology, mixed with the model of deliberation based on Foot's exposition of external reasons, neither how one comes to know an external reason, nor how one

Stephenson

comes to know how that reason suggests an agent should act must lead to coercion. Thus, features of the morality system which make external reasons seem inescapable do not apply. Because, according to the model I set out above, external reasons do not clear easily knowable dictates for action, external reasons can be seen to factor into deliberation alongside internal reasons. An advisor cannot blame an agent for simply failing to act on an external reason without engaging the agent's S and trying to understand the situation. The version of external reasons I have sketched above requires that the advisor engages the agent's S before even determining what the agent should do, so this scenario does not appear to be very different than where Williams ended up with normative internalism. In both cases, the interlocutors can reach an impasse in the argument. Once the argument reaches this impasse, the advisor seems no more or less likely in either strong normative internalism, or a reasonable conception of externalism, to engage in coercion as the next step in getting the agent to act. Williams, then, has to either maintain weak normative internalism, which does not fit with his attempts to make internal reasons normative, or concede that his strong normative internalism is no more or less likely to lead to coercion than external reasons.

Conclusion

Although debates about the nature of reasons hold important philosophical merit and perhaps even useful practical insight, we must not deceive ourselves into thinking abstract philosophical positions have more dire consequences than they, in fact, do. Williams, unfortunately, fell into this sensationalist philosopher's conceit when claiming there was great danger of coercion in external reasons and that his normative internalism stood out against external reasons as bearing some special protection against unnecessary coercion.

By allowing that an agent can be irreconcilably mistaken about what she has reason to do, Williams opens internalism up to the possibility of coercion. Because coercion goes against what an agent wants, not necessarily against her S, the possibility is open that the advisor can coerce the agent into

Stephenson

acting according to her S, but against her wishes. Williams' strong normative internalism turns out to offer no special protection against such coercion.

Not only does Williams exaggerate the un-coercive nature of his normative internal reasons, he also mischaracterizes the nature of external reasons. Just as nothing about the nature of normative internal reasons suggests that they hold any special protection from coercion, nothing about the nature of external reasons suggests that they are necessarily coercive. Williams needlessly ties external reasons to the special kind of obligation he sees in the morality system, leading him to erroneously claim that appeals to external reasons, failing rational persuasion, require coercion. Simply by separating external reasons from the morality system and revealing how they can be integrated into deliberation alongside other conditions, I have disentangled external reasons from coercion. The primary content of reasons has little to say about how an agent should act and almost nothing to say about how an advisor might go about persuading that agent to act.

None of this suggests much about when coercion is justified. I have merely pointed out that Williams sensationalized the danger in external reasons. That being said, there is great value in the way Williams goes about setting up normative internal reasons and criticizing what he thinks external reasons have to be. Fortunately, however, nothing suggests that this disconnected picture of rational objectivity is the right way to conceive of reasons that get their force from considerations external to our subjective leanings.

This is an interesting result because it suggests that once we start to really scrutinize reasons for action, whether they are internal or external actually seems to have little ethical significance. What really seems to matter in persuasion is not exactly what an advisor appeals to when trying to get an agent to act, but that we make a genuine attempt at understanding the agent's perspective. While this does not settle the question of how viable the concept of external reasons is, it shows that external reasons do not carry a

special tie with coercion. Lacking any special tie between external reasons and coercion, we need not discourage appeals to external reasons based on fears of coercion.

Bibliography

- Chapman, John W. and Penncock, J. Roland ed. *Coercion*. Chicago: Aldine, 1972.
- Foot, Philippa. *Natural Goodness*. Oxford University Press: Oxford, 2001.
- McDowell, John. "Might there Be External Reasons?", in J.E.J. Altham and Ross Harrison ed. *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge University Press: Cambridge, 1995.
- Louden, Robert. "The Critique of the Morality System" in, Alan Thomas ed. *Bernard Williams*. Cambridge University Press: Cambridge. 2007.
- Scanlon, T.M. *What We Owe to Each other*. The Belknap Press of Harvard University Press: Cambridge, 1998.
- Williams, Bernard. "Internal and External Reasons", in *Moral Luck*. Cambridge University Press: Cambridge, 1981.
- Williams, Bernard, *Ethics and the Limits of Philosophy*. Harvard University Press: Cambridge, 1985.
- Williams, Bernard. "Internal Reasons and the Obscurity of Blame", in *Making Sense of Humanity*. Cambridge University Press: Cambridge, 1995.
- Williams, Bernard. "How Free Does the Will Need to Be?", *Making Sense of Humanity*. Cambridge University Press: Cambridge, 1995a.
- Williams, Bernard. "Postscript" in Elijah Millgram ed. *Varieties of Practical Reasoning*. MIT press: Cambridge, 2001.
- Williams, Bernard, "Values, Reason and The Theory of Persuasion", in A.W. Moore ed. *Philosophy as a Humanistic Discipline*. Princeton University Press: Princeton, 2008.