

REASONS FOR ACTION

Anthony Tiberio

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy.

Chapel Hill
2010

Approved by:

Susan Wolf

Ryan Preston-Roedder

Robert Merrihew Adams

© 2010
Anthony Tiberio
ALL RIGHTS RESERVED

ABSTRACT

Anthony Tiberio: Reasons for Action
(Under the direction of Susan Wolf)

What is the correct relation between an agent's motivations for action and his normative reasons for action? Those who think that there is some necessary link between motivating and normative reasons are those who accept an internalist view about normative reasons. Those who reject this idea hold an externalist view about normative reasons. In this paper, I will present, analyze, and address two challenges levied against externalism from internalist perspectives. The result of my analysis will hopefully demonstrate that the first challenge against externalism does not refute externalism, and that there is not sufficient justification for adopting one notion of rationality over another. As a result, part of what is often referred to as the internalism and externalism debate appears to be merely a terminological dispute.

TABLE OF CONTENTS

Section	Page
1. Introduction.....	1
2. What Are External Reasons?	6
3. What Is the Point of External Reasons?.....	14
4. Beyond Pragmatics	18
References.....	25

1. Introduction

What do I have (most) reason to do? I can pose this question looking for different sorts of answers. I might ask this question with an eye to figuring out what my long-term projects should be or should involve. I might raise this question while trying to figure out what I should do right now. Either way, if I am asking someone this question, it is often natural for the respondent to base her reply upon what it is that I care about.

However, is this correct? Are facts about what I care about the only relevant facts in determining the correct answer here? Are such facts even necessary to determine what I have most reason to do? What is the relationship between what I have reason to do and my motivations?

Must an agent have a certain motivational state in order to have a reason to do X? When employing one sense of ‘reason,’ this would be true by definition. We can call this sense of reason, a *motivating reason*. A motivating reason for an agent is a desire to perform a particular action – a desire that an agent has, and based on which he or she acts. This kind of reason can serve to explain why an agent performs a particular action. That is, a motivating reason can appropriately help to answer the question “why did the agent perform this action?” A second sense of ‘reason’ is an *explanatory reason*. Explanatory reasons include more than motivating reasons. Explanatory reasons include other causally relevant factors behind an agent’s action. For example, when describing the situation where a patient’s leg kicks up when struck on the knee with the doctor’s mallet during an examination, we can say that ‘the reason he kicked was that the doctor hit his knee.’ This use of ‘reason’ serves to

explain why the agent's leg moved, but this reason is neither a normative reason nor a motivating reason.

The third kind of 'reason' is a *normative* reason. Normative reasons are considerations that count in favor of an agent performing some action. Normative reasons determine whether or not an agent's motivating reason on a particular occasion was a good or bad reason for acting. So, when asking whether or not an agent must have a particular motivational state in order to have a reason for acting, we are inquiring into the nature of normative reasons.

This thesis will focus on questions regarding the correct relation between motivations and normative reasons. Those who agree that there is some necessary link between motivating and normative reasons are those who accept an *internalist* view about normative reasons. That is, the internalist asserts that whether or not an agent has a reason for acting is always dependent upon facts about the agent's motivational state.

Those who reject this idea hold an *externalist* view about normative reasons. Externalism about normative reasons claims that facts about an agent's motivational state are not the only facts relevant to determining what that agent has reason to do. Such facts about an agent's motivational state may not even be necessary in at least some cases. In other words, externalism asserts that on at least some occasions an agent will have a reason to do a particular action no matter what the agent's motivations happen to be.

A more precise formulation of this distinction is as follows. Internalism maintains that a necessary condition for R is M where:

(R) Agent A has a reason to X

(M) A is motivated to do X when A has the relevantly true beliefs; or, A is motivated to do Y (where X is, at least, a way of helping to bring Y about) when A has the relevantly true beliefs.

Following Bernard Williams, we can use the following case as our guide to understanding internalism in this way (Williams 1982: 102). Imagine an agent who is sitting at a table with half-full glass of tonic water and a bottle marked 'gin.' This agent has a strong desire to have a gin and tonic. Unfortunately, the 'gin' bottle does not contain gin, but contains petrol instead. Does this agent have a reason to X, where 'X' is to mix the stuff in the gin bottle with his tonic water and drink it? Of course, if the bottle actually contained gin in the bottle, then the agent would have reason to do X. But given his false belief (that 'that stuff' is gin), we do not want to say that he has reason to drink what's in the gin bottle *even though* he has a desire to drink what's in the gin bottle.

The internalist, qua internalist, is committed *only* to the claim that M is a necessary condition for R. It is open to the internalist to claim that M is also a sufficient condition for R. Let us call the view that claims that M is a necessary condition, but not a sufficient condition, for R *weak internalism*. Let us call the view that claims that M is both a necessary and sufficient condition for R *strong internalism*.

It is unclear what sort of internalist Williams is, but it is worth asking what the motivation is for being one and not the other. On one hand, one might be motivated to be a weak internalist if one does not want to be committed to the view that every urge, whim, and compulsion, provides a reason for action; perhaps only desires and cares that the agent has and identifies with are sufficient to generate reasons for action. But on the other hand, while it might seem odd to think that even our faintest whims can provide a reason for an action, it is not implausible to think that as long as you have the relevantly true beliefs, your whims or

fleeting desires, at that time, provide *some* consideration in favor of your doing X. If the presence of some fleeting desire counts as *some* consideration in favor of doing X, however small, then it is a reason for doing X.

In contrast to these forms of internalism, an externalist view is any view that claims that M is *not* a necessary condition for R. That is, an externalist view claims that there are some reasons for action that are determined by factors *other than* the motivational states of A. The externalist has in mind the possibility of normative factors being legitimate reasons for actions, even if such normative factors are not part of (or relevantly linked to) the motivation of the otherwise morally capable agent.

There are two ways to be an externalist. *Strong externalism* is the view that any fact about an agent's motivational states is irrelevant to whether or not the agent has reasons for action(s). On this view, all reasons for action are determined completely independently of the contingent nature of an agent's motivational makeup. According to *weak externalism*, some reasons for action exist only if an agent possesses certain motivational states while other reasons for action exist independently of M.

A second distinctive feature of externalism, as pointed out by Derek Parfit, is that if R, then E (Parfit 1997: 101) where,

(R) A has a reason to X,

(E) If A knew all the relevant facts and A is substantively rational (i.e. knew the relevant facts and *cared* about the appropriate objects of value), then A would be motivated to do X.

According to Externalism, E is simply a consequence of R, and R's truth does not depend upon M. If R is true, M could also be true, but R can be made true by factors other than M. Thus, R and M are logically independent of one another, in at least some cases.

In this paper, I will present, analyze, and address two challenges levied against externalism from internalist perspectives. The first challenge to externalism is that it is not clear what it even means to have an external reason for acting. This challenge can be understood in a few different ways, but I will show that the best way to understand it is as a skepticism about there actually being certain sorts of normative facts. I will briefly address this worry and then show that this challenge to externalism is only one horn of the internalism/externalism debate. It is a challenge that can be put forth against any objectivist theories of normativity.

The second challenge to externalism brings out the second horn of the internalism/externalism debate. The challenge is that if externalism is true, there seems to be no point in claiming that A has a reason to X so long as externalism is true. I try to first assuage this worry by pointing out the pragmatics of normative discourse; but, it appears that the underlying motivation behind this challenge is the internalist's dissatisfaction with the externalist's substantive conception of rationality. I then offer some preferences in favor of adopting a substantive conception of rationality propounded by the externalist. But I conclude that there is not sufficient justification for adopting one notion of rationality over the other. As a result, the crux of the second horn of the internalism/externalism debate appears to be merely a terminological dispute.

2. What Are External Reasons?

What does it mean to say that someone has a reason to do something, *even if his or her actual motivation is in no way tied to this action*? Williams says that he finds it ‘quite obscure’ what this could mean (Williams, June 1995: 40).

There seem to be two ways in which someone could be unsure of how to understand what it is for an agent to have an external reason for acting. One way is to understand proposition R (where R: ‘A has a reason to do X’) as simply *meaning* M. We can call this position *analytic internalism*. If R semantically contains M, then it would be quite difficult to understand what the externalist is asserting, and it would be natural to think that the externalist simply fails to adequately understand what R means. R just meaning M would be an example of the strong internalist position, since M would be sufficient for R, not only necessary.

But why think that R is either semantically equivalent to, or just semantically contains, M? Ordinary usage of normative statements, of the same form as R, in everyday life does not indicate either of these intimate connections. In fact, many ordinary uses suggest the opposite. Take someone who while at a party mistakenly walks into an empty guest room and sees a stack of cash on the floor. Realizing that the chance of getting caught is slim (perhaps because no one is around at that particular time, and that a large number of people has been near the room all night), he takes the money. We might say “The reason he stole the money was that the money was unattended and out in the open.” This is a use of reason used solely in the explanatory, non-normative and non-motivating, sense.

But in this same case we can also say “The reason he took the money was because he wanted it and realized that taking it wouldn’t bring him shame from other people, shame which he desired to avoid.” This use of reason not only explains why the agent took the money but does so by specifying the agent’s motivating reasons for taking the money.

A third thing we could also say about the same case is: “The fact that the house owner’s money was out in the open while no one was around gives him no reason to take it.” This use of reason here is clearly of the normative type. Such a usage implies that simply wanting the money and not wanting to get caught do not suffice to form *good* reasons for taking the money. Even if the internalist is right, and this externalist normative claim were false, it is certainly an ordinary linguistic use of ‘reason.’

Given this, the default position cannot be to understand R as simply meaning M since ordinary usage conflicts with this being the case. Further evidence is needed to support analytic internalism. I doubt there is such evidence.

One could also fail to see what the externalist means by R in a way that does not presuppose that R semantically entails M. We can call this position *non-analytic internalism*. This sort of internalist maintains that R and M are merely contingently identical to one another. That is, in this world, as it turns out, facts about what reasons one has for action just happen to be facts about how an agent is motivationally constituted. But, this need not be the case in all possible worlds. It just so happens that in our world the proposition, R, is true when, and only when, the proposition, M, is true. Thus, an internalist can maintain that all normative reason statements about an agent hold only when certain propositions concerning facts about the agent’s motivational states hold, without thereby being committed to claiming that each normative reason proposition (that is, each R) is semantically equivalent to some

proposition about the agent's motivational states (that is, some M).

In turn, one could be a strong non-analytic internalist or a weak non-analytic internalist. According to the weak internalist, wherever R holds, M holds; and, according to the strong internalist, wherever R holds, M holds and wherever M holds, R holds.

Non-analytic internalists (either strong or weak) claim that they are unclear regarding what R means when understood in the externalist way. What are they unclear about? Let us consider the following.

For non-analytic internalists, the class of actions that an agent has reason to do is equivalent to or is a subset of the class of things for which M would be true. The externalist asserts, instead, that some actions that an agent has reason to do *fail* to have property M.

The worry that the internalist has is this: why think that the class of actions that an agent has reason to do contains the actions that the externalist claims it includes? The internalist has a story to tell regarding why R refers to the class it does --- namely, because M is present. The externalist offers no such story; no necessary or sufficient conditions are offered. A set of necessary or sufficient conditions must be given if the account is not to be 'obscure,' (as Williams finds it).

Derek Parfit tries to offer some such story. He writes,

"Reasons for acting, we might say, are facts that *count in favour* of some act. But 'counting in favour of' means 'giving a reason for'. Or we might say that, if we have *most reason* to act in some way, that is what we *ought rationally* to do, or—more colloquially—what we *should* do. But we could not understand this use of 'should' unless we had the concept of a reason," (Parfit 1997: 121).

He continues,

"These two concepts—that of a normative reason, and the concept that is expressed by this use of 'should'—cannot I believe be helpfully explained, since they cannot be explained in non-normative terms. This fact is not surprising. Normative concepts form a fundamental category—like, say, temporal or logical concepts. We should not

expect to explain time, or logic, in non-temporal or non-logical terms. Similarly, normative truths are of a distinctive kind, which we should not expect to be like ordinary, empirical truths. Nor should we expect our knowledge of such truths, if we have any, to be like our knowledge of the world around us,” (Parfit 1997: 121).

Here, Parfit is offering something akin to a definition of what a reason for action is. A reason for action is a fact that counts in favor of that action. However, Parfit claims that this concept of a fact ‘counting in favor’ of some action cannot be further analyzed into non-normative terms. Parfit thinks that these normative facts cannot be ‘helpfully explained.’

Now, just by understanding a reason for action as a consideration in favor of doing that action, we can reconstruct the internalist’s worry. Why are externalist reasons for action obscure? They are obscure because we want to know which facts are the ones that count in favor of doing some action. If these facts are not the M facts, facts about the agent’s actual motivations, then what could such favorable or unfavorable facts be?

If an externalist were to end his explanation as Parfit does above, then the internalist’s worries would have some merit. But, fortunately, I think that more of a story can be told by the externalist; a story that Parfit does not tell. Some facts that are relevant in determining reasons for action are facts that have an impact on whether or not: (i) the world is made better or worse; (ii) someone’s rights are violated; (iii) the agent is made better or worse off; (iv) the agent’s desires are satisfied.

The internalist might press further: why think that these are the relevant facts in determining what it is that an agent has reason to do? These facts are only relevant (according to the internalist) if the agent cares about his welfare and morality. In response, I think that the externalist should simply maintain that there is no explanation as to why these factors are relevant in determining what an agent has reason to do. I agree with Parfit in

thinking that, at some level, the externalist needs to admit that there is no helpful explanation to be had when trying to explain why certain facts are relevant to determining what normative reasons an agent has (Parfit 1997: 109). I disagree with Parfit on precisely where the externalist needs to maintain that the explanation stops.¹ And if there is no further explanation to be had here, as the externalist maintains, this could be because the categories above are (or reflect) primitive normative facts. That is, there is no way to explain these normative facts in terms or in relation to facts about particular agents' motivational states.

We could understand the externalist as saying that there is no way to *reduce* these normative facts to facts concerning agents' actual psychology. Put in this way, I don't think it follows that an externalist, qua externalist, *has* to be a non-reductivist in general, as Parfit argues (Parfit 1997: 129). However, an externalist is committed to the claim that facts about normative reasons cannot *solely* be reduced to (or derived from) facts about M. I am not sure what sort of facts these normative facts could be reduced to or derived from, if any. Still, I do not wish, *prima facie*, to rule out the possibility for a naturalistic reductive account of normative facts. Yet, it does appear that the onus is on the optimist about such a reductive project to show that it can be done (where a naturalistic view is one where normative facts are reduced to natural facts – facts which are presumably describable in terms of the natural sciences).

Parfit does not say much about the nature of the normative facts that ground our normative reasons. Parfit writes, "...reasons for having the desires on which we act...are reasons to want *some thing*, for its own sake, which are provided by facts about this thing.

¹ Parfit does not actually tell the externalist story as I do but from what positive externalist view can be gathered from what he says in "Reasons and Motivation," my story may even be compatible with his view.

Such reasons we call *value-based*,” (Parfit 1997: 128). He adds that “such alleged normative truths may seem to be metaphysically mysterious, or inconsistent with a scientific world view” (Parfit 1997: 129). This mysteriousness does not dissuade Parfit from thinking that there are such brute, normative facts. He thinks that one of the worlds that has these sorts of un-analyzable normative properties happens to be our world.

An internalist might doubt that this is the case. This doubt can be expressed by questioning what it even means to call one state of affairs ‘better’ than another. But I don’t know that much can be said in response to this worry. If being ‘better’ is a primitive normative concept, no explanation can be given. Perhaps, all we can go on is our intuitive judgments made in particular cases to guide us into deciding which world is better than another. Such judgments might be what guide us in determining what is the better outcome (and in turn, what we have reason to do).

Of course, it is a difficult task for the externalist to sort out when the above factors (i through iv) are relevant and how much weight each should be given. These are questions about which different externalists will disagree. They may not be ones directly relevant to our current discussion. They are relevant in determining what one has most reason to do within externalist normative schemes. The current question is whether externalism has illuminating things to say about the idea of having reasons (or good reasons) for action. The issue of whether an agent’s rights or best consequences of an action determine the best normative reason for an action, for instance, does not seem to be directly pertinent to the internalism-externalism debate.

On the other hand, might it be that what underlies Williams’ obscurity charge against the notion of external reasons is precisely the fact that there is so much disagreement among

externalists regarding what is better, what it means to have well-being or rights, etc. and, thus, what it is that we have most reason to do? He might think that these disagreements indicate that there are no real normative primitives of the sort Parfit claims. This is a substantial worry that I am unable to fully address here. But the story the externalist needs to tell should take the following form. Since we are dealing with primitives, accessible perhaps only by intuition and often in the face of impaired or culturally tainted intuitions, such disagreements are to be expected.

Also, it should be pointed out that this sort of challenge to externalism makes sense only when made by a certain type of internalist. This is the sort of internalist who is skeptical that categories (i), (ii) and perhaps even (iii) are ultimately coherent. Let us call this sort of internalism *skeptical internalism*. (Obviously, one can be a skeptical weak internalist or a skeptical strong internalist).

Skeptical internalism is to be contrasted with *objective internalism*. Objective internalism would claim that there are certain states of affairs or traits/characteristics of agents (and of the world) that are objectively good/bad, but R can only be true when M is true. (Again, one can be either an objective weak internalist or an objective strong internalist).

So, when our internalist complains that external reasons are obscure, this internalist might be resisting a non-naturalistic account of normativity. I do not think that simply because an explanation is non-naturalistic it should be rejected. If the explanation accounts for its putative explanandum, so much for naturalism.

Of course, much more needs to be addressed and analyzed in order to adequately deal with understanding the foundation (or lack thereof) of normativity.

What I hope to have shown here is that the skeptical internalist complaint about the obscurity of external normative reasons, is a complaint at a metaethical level of debate. It is a complaint about the nature and possible reality of values. Or put differently, it is a worry that categories (i), (ii) and (iii) ultimately make no sense. The debate over whether or not these categories are tenable or are ultimately coherent is not a debate that needs to be dealt with in the context of the internalism-externalism debate. In fact, it is somewhat misleading to characterize this debate (the debate about objective values) as internalism versus externalism, since the objective internalist is on the side of the externalist in defending the coherency and objectivity of normative categories (i) through (iii).

But it is important to understand that this debate about the foundation of normativity is very relevant to another debate that falls under the title of ‘internalism versus externalism’. This other debate is a debate about how we should understand what it is to be a rational agent. Is it a requirement of rationality that an agent *care* about certain things? Importantly, this question seems to arise only if one takes it for granted that there is some sense in which it is better to care about certain things. Thus, the two debates are related.

3. What Is the Point of External Reasons?

Consider a variation of the infamous shallow pond case. Fred is running a bit late while on his way to the airport in order to depart for a holiday destination. It was a fairly inexpensive plane ticket. While walking to the airport, Fred notices a child struggling to stay afloat in a small pond. Fred could swim out and save the child, but since it will take some time and he is running late for his flight, he will miss his flight. Noticing this fact, Fred decides to keep going and to try to make his flight. Moreover, Fred lacks the desire or motivation to be good (in the *de dicto* sense). That is, Fred lacks a desire to do what is good, *qua* being good.²

Let us assume that M is false in this case. Fred is not motivated to save the child, nor is he motivated to do something else that would be helped or served if he were to save the child. An externalist (or at least a plausible kind of externalist) will claim that despite M not being the case, R still holds. Fred has good reasons for saving the child (because the child's death is a much worse thing than are Fred's inconveniences should he save the child).

Bernard Williams, with a case of the same sort in mind, writes,

“it would make a difference to ethics if certain kinds of *internal* reasons were very generally to hand... But what difference would external reasons make? ... Should we suppose that, if genuine external reasons were to be had, morality might get some leverage on a squeamish Jim or priggish George, or even on the fanatical Nazi? ... I cannot see what leverage it would secure: what would these external reasons do to

² Just because Fred lacks a desire to do what is good, *qua* being good, does not mean that Fred does not care about anything that is good. He may care about good things; but, he does not simultaneously (a) care about a good thing; (b) know that it is good; and, (c) care about it because it is good.

these people, or for our relations to them?” (Williams, April 1995: 215).

Williams is claiming that if internalism is true, we can make sense of what is gained by saying that you have a reason to do X. Whereas if externalism is true, there is nothing gained by this claim that is not captured by saying, in the case of Fred, for example, that ‘X is what a decent person would do.’ This argument by Williams assumes that for the reason statement to ‘gain something’ (or to ‘make a difference’) normative reasons need to have a sort of *leverage* on people. The sort of leverage that Williams has in mind seems to be the kind that would convince them to be motivated differently.

Parfit directly addresses this argument by Williams by accusing Williams of having a conception of ethics that is “too utilitarian.” Parfit thinks that claiming that some agents have reasons for acting is not a way to affect those agents’ ways of acting. Rather, the aim of holding such beliefs is “not [about] influence, but truth,” (Williams, April 1995: 215). Parfit thinks that Williams makes the mistaken assumption that the truth of normative claims necessarily involves/entails that normative claims are aimed at motivating people to act in certain ways, rather than to ‘get at the truth’. I think Parfit is right to think that the truth of a normative statement does not depend upon one’s motivational makeup; but, Parfit is wrong to think that we make normative claims solely with the aim of ‘getting at the truth.’ In fact, we utter normative claims often with the *sole* purpose of trying to motivate ourselves and others to conform to the content of the asserted norms.

As the internalist suspects, the leverage that we associate with normative claims does vanish, in a sense, if the agent in question cannot become motivated to do what is good and decent. It does not seem that leverage or ‘normative force’ is required for an external reason statement to be true, but leverage or normative force is often sought when making utterances

about what particular agents have reason to do.

In cases where M does not hold for A and motivationally cannot so hold, it seems futile for one to tell A that he has a reason to X. If an utterance of this sort is absolutely futile, this counts against an agent's making the utterance to A. Telling Fred that it is more important to save the child than to catch the plane is pointless if Fred does not care about children or does not care about what is important, and the speaker knows this about Fred.

However, importantly, this is not evidence that A does not actually have a reason to do X. What this does show, however, is that informing an agent about what this agent has reason to do is *not* determined solely by whether or not the agent has good reason for action X. It is also dependent upon whether or not there is any chance that the agent will be convinced to come to want to do X. That is, it is also dependent upon whether or not the agent actually *can* become motivated to do X.

This is an important point for the externalist to underscore. Of course, the internalist is going to doubt that there is an 'actual' external reason for action when M does not hold for the agent in question. If the internalist's worry here is that if we understand normative reasons as external reasons, we would be *completely* divorcing normative claims from motivating people to be decent or prudent, this is simply not the case. It is true that if an agent has reason to do X, but cannot ever become so motivated, it *may* be inappropriate to *utter*, to the agent, that the agent has such a reason. If making the claim that 'A has reason to do X' will not have any affect on this person's behavior, this counts against one uttering this claim to the callous agent. We could still make the claim in a general way, as in making a list of all true statements about our world.

An externalist, perhaps Parfit, might disagree with placing so much emphasis on

motivating people to do what they have reason to do. In everyday life, many people, including many philosophers, constantly make claims of the sort: ‘the president has reason to do Y’ and ‘my friend has reason to do Z’ and often times we make this claims *without even knowing*, or caring, about what the agent’s actual motivations are. Even if the president has no intention to decrease the size of the inflated US defense budget in order to help assuage human rights violations (nor will he develop one), we might still be justified in claiming that he has reason to do so.

While it is true that we might be justified in making normative claims about what the president and our friend has reason to do (without knowing whether or not M holds), this is not sufficient evidence to show that we do not usually utter normative statements with the intent of motivating *some* agents. I doubt one would tell the president that he has reasons to decrease military spending if one truly believed that the president would never become so motivated. In actual situations it may not be possible to be certain about Obama’s future motivations. We may tell him that he has reason to decrease military spending with the hope of planting a seed that might change his motivation in the future.

Making normative utterances with the purpose of changing people’s motivations is more salient in cases where one speaks to a friend, colleague, relative, etc. Uttering normative claims about other agents often does play a role in influencing the interlocutor’s cares and desires. We often are trying to *give or show* our interlocutor what we think are the correct, or most salient, reasons for action. Furthermore, uttering such normative claims also plays a role for the speaker to identify, evaluate, and amend or reaffirm what the speaker cares about. At the very least, it serves to create (or reaffirm) the following motivation: ‘if it was within my power, I would have reason to drastically decrease military spending.’

4. Beyond Pragmatics

One may agree with the pragmatic emphasis placed upon affecting others' and one's own motivation when engaging in normative discourse, but still reiterate Williams' challenge.

Let us grant that the main purpose of uttering R is to influence others and ourselves (where who is being influenced depends on the context of the utterance). An internalist may challenge that, presumably, the goal of uttering either (B) 'It would be better if A does X' or (C) 'It would be better for A if A does X,' is different from the goal of uttering (R) 'A has reason to do X'. Presumably the externalist wants to maintain that there is *something* extra gained by R than is gained by the former two claims. What is different about these? What extra 'normative force' does R have that others lack?

I think that the externalist should simply deny that (R) has any more 'normative force' than (B) or (C). It is a mistake to think that something extra is had by one and not the others. To claim that 'an agent has a reason to do X' is just to claim that 'it would be better if the agent did A.' If it is a prudential case, it is better *for the agent*. If it is a moral case, it is *morally better* that A does X.

An internalist may be dissatisfied with this consequence of externalism. Williams writes,

“what is the difference supposed to be between saying that the agent has a reason to act more considerately, and saying one of the many other things we can say to people whose behaviour does not accord with what we think it should be? As, for instance, that it would be better if they acted otherwise?” (Williams, June 1995: 39-40).

As I've contended above, there need not be a significant difference between the two sorts of claims. But I suspect that what bothers Williams is that the externalist is expanding the

notion of rationality without sufficient justification for doing so. Williams seems to imply that if we were to gain something extra (perhaps some extra motivational force) with claim R, then the externalist would be justified in stretching the notion of rationality. But since this is not the case, we should restrict our notion of rationality to the internalist sense.

Consider the shallow pond case. The externalist wants to claim that Fred both has a reason to save the child and a reason to catch his flight on time, but since the former is a stronger reason than the latter, he has most reason to save the child. That is, he ought to save the child. Since if he fails to save the child he acts contrary to what he has most reason to do, Fred is acting irrationally.

An internalist who claims that it is actually worse if the child drowns than if Fred misses his flight is an *objective internalist*. This sort of internalist is to be contrasted with a *subjective internalist*. The subjective internalist is the sort of internalist who is skeptical that normative reason statements, when used in an external sense, have any semantic content or any actual referent in our world.

Let's compare the objective internalist's description of the shallow pond case with the sort of externalist description that I have given above. We are tabling the subjective internalist position here because it is the objective internalist position that most starkly brings out the issue of concern here: what is it to be rational? By focusing on subjective, instead of objective internalism, we run the risk of conflating two issues into one: the foundation of normative claims and what it is to be rational. These issues are related in that the question of what it is to be rational is not a controversial matter if there are no objective values: it's rational to pursue what you care about, more or less. If there are objective values, it matters, in some sense, what you care about. Given this, the question of what is required of an agent

to be rational, becomes a more interesting question. Does being rational require that an agent care about certain things?

The objective internalist wants to simultaneously maintain both that it is worse if the child dies than if Fred misses his flight, and that Fred as actually constituted lacks a reason to save the child. According to the internalist, since Fred lacks a reason to save the child (perhaps because Fred dislikes children), he is not irrational in leaving the child to drown. In fact, as the case is stipulated it appears that it is actually *more* rational if Fred leaves to catch his flight, which is important to him, rather than attend to the child.

Thus, we are left with the externalist and the objective internalist agreeing regarding which state of affairs is better and which is worse, but disagreeing regarding whether or not to call Fred's behavior irrational. So, are we left with merely a semantic debate?

On the surface, it is a semantic debate. But, I think that the externalist conception of rationality may actually cohere better with other ordinary normative judgments that we make. The externalist wants to understand a rational agent as more than an agent who has the relevantly true empirical beliefs and who acts in accord with what she cares. The externalist understands a rational agent as being one who also *cares about the right things*. This is a substantive notion of rationality.

The internalist wants to restrict the notion of rationality to the strictly procedural aspect. This version of rationality can be understood as being akin to the notion of a mathematical function. Start with the inputs (i.e. your motivational states when having the relevantly true beliefs), and the outputs will come out in accord with the content of these inputs. The more one acts in accord with these inputs, the more rational one is. The less one acts in accord with these inputs, the more irrational one is. The externalist takes this notion and expands it.

The externalist agrees that this function is part of what's required to be rational, but having (or lacking) certain inputs also determines whether or not one is rational. Given Fred's immoral behavior, some might even call him crazy or a psychopath. Is this claim substantially different than calling Fred irrational? I fail to see that it is.

Of course, the objective internalist is going to resist this. The objective internalist will point out that there is not sufficient justification for building facts about what an agent ought to value into our concept of rationality. Sure, it is objectively true that the moral and prudential person possesses some traits and lacks others, but this does not mean that this person is *rational*. Though the moral person, or the prudential person, possesses many good and perhaps even desirable traits, each can still be utterly irrational. Why not just keep normative concepts about reasons (like rationality or irrationality) separate from normative concepts concerning values or virtues?

Ultimately, I do not think that there is sufficient reason to adopt a substantive notion of rationality over a procedural one, nor is there sufficient reason to adopt a procedural notion of rationality over substantive one. Yet, I prefer to adopt of substantive notion of rationality because we are able to assert the following sorts of claims that I find highly intuitive. In the shallow pond case, the salient facts are that (F) 'the child will drown if Fred doesn't help' and that (F') Fred's helping the child will only cause minor inconveniences. Action X is that Fred jumps in the water and saves the child. From facts F and F', it seems to follow (however indirectly) that it is morally wrong for Fred not to X. If it is morally wrong for Fred to have not done X, then, morally, Fred ought to have done X.

These claims about Fred thus far seem to be innocuous to both the internalist and externalist. But it is the next (externalist) step(s) that I find highly intuitive. From the fact

that Fred morally ought to have done X, it seems to follow that Fred ignored important moral consideration(s) for acting. Since he ignored important considerations for acting, he seems to be acting irrationally because he, as an agent, is not properly cohering with what is valuable. Or rather, he is failing to properly pick up on something – namely, what is objectively valuable. This phenomenon of failing to do what is valuable or worthwhile, as in Fred's example, is a phenomenon where an agent is disconnected with the world in some important way. The internalist's notion of procedural rationality allows only for a disconnect within the agent as irrational – one where the agent's desires/cares don't properly match up with how the agent acts. The reason I prefer to use a substantive notion of irrationality is that it permits us to categorize both the internal and external disconnects that an agent might have as the same sort of phenomenon.

The objective internalist may respond by making a point that Williams does in this context. Williams says that some philosophers take 'A has most reason to X' to be equivalent to 'A ought to do X.' Williams adds that he does not want to presuppose their equivalence (Williams 1982: 106). I *do* want to assert their equivalence. This allows us to categorize two different phenomena as being of the same *kind*. I think that the most natural way to understand what it is that one ought to do is in terms of what one has most reason to do. Fred is indeed being both morally callous and irrational. At the very least, it seems as though Fred is not being fully rational – though he is being procedurally rational.

The objective internalist appears to have two different alternatives if he wants to remain an objective internalist: (1) try to tell more of a story as to why these claims are not in fact equivalent (i.e. the internalist should further elaborate on why we should think that one can be rational without doing what one ought to do); (2) deny that having an important moral

consideration to do X entails that one ought to do X. Or, if one wants to give up internalism, one can maintain the following claim that I find very intuitive: the amoral person is actually irrational since she fails to do what she ought to do (and in turn, what she has most reason to do).

The first option seems implausible. Saying that A ought to X seems to only make sense if A has a reason to X. The second option seems implausible as well. If one fails to recognize a morally compelling reason for acting, isn't this a sort of failure on the agent's part that can appropriately be called irrational? We do call Hitler and Stalin crazy. Why not call them irrational? These agents are responding inappropriately in thought, feeling and action to the considerations that *should* serve as motivating. If one insists on calling Fred rational (and on calling Hitler rational) there may not be much that can be said in response. We would have reached a stalemate.

Ultimately, the internalism and externalism debate is simply a verbal dispute. It is a verbal dispute over what to call agents like Fred or Hitler. Hitler was certainly rational in the procedural sense, but was crazy in another sense. This 'craziness' is captured by the externalist's substantive notion of rationality, while the former procedural sense is captured by the internalist's sense of rationality. My preferring the substantive notion of rationality could just be my preference for a certain terminology (and thus, for a certain sort of taxonomy). Ordinary usage of 'reason' and 'rationality' talk seems to allow for both conceptions, and, as a result, cannot help to adjudicate the matter.

But even if we were to decide to use rationality in the procedural sense, both the externalist (as depicted in this section) and the objective internalist views share part of the same spirit. The deeper divide is between the skeptical internalists and objective internalists

(and externalists). The skeptical internalist denies the objective status of normative claims. But this debate is the standard debate between skepticism and objectivity in value-theory.

This is probably where the action really is, but that action falls outside the purview of this paper.

References

- Parfit, D. 1997. Reason and Motivation. In *Proceedings of the Aristotelian Society*, Supplementary Volume 71, 99-130.
- Williams, B. 1982. *Moral Luck*. Cambridge: Cambridge University Press.
- Williams, B. April 1995. Replies. In *World, Mind, and Ethics*, eds. J.E.J. Altham and R. Harrison. Cambridge: Cambridge University Press.
- Williams, B. June 1995. *Making Sense of Humanity*. Cambridge: Cambridge University Press.