

# REASONS AND RATIONALITY

Sven Nyholm

The questions that I will be discussing are:

Q1: What is the relation (and difference) between normative reasons and motivating (and/or explanatory) reasons?

and

Q2: When are, or in virtue of what are, acts, desires or beliefs rational (or irrational)?

Because the two subjects that I will discuss—reasons and rationality—are so closely related, I don't think that one can discuss one without mentioning the other. I shall, however, have more to say about Q1. Questions about reasons are, I believe, more fundamental than questions about rationality.

Following Bernard Williams<sup>1</sup>, I distinguish between internalism and externalism about normative reasons. Internalism is the view that there is a necessary connection between our normative reasons and our motivations. Unlike Williams I believe that we should take the externalist view, which is the view that there is no such necessary connection.

One of Williams' claims is that claims about external reasons make little or no sense. I argue that, though they cannot, by means of reduction, be properly or thoroughly explained, such claims do make sense if we take them to be irreducibly normative. Thus we don't need to explain what it is to believe that one has some reason; what one believes is simply that one has this reason.

Reasons are, on the view that I will be defending, facts that count in favour of caring about certain things or living in certain ways. A fact is a reason if it has normative significance and if it, thus, matters. However, to say that something matters is basically to say that there is reason to care about it. Thus claims about reasons cannot, just like I said above, be reduced to claims about other things that aren't normative. Claims or beliefs about normative reasons are, if they are to be truly normative, irreducibly normative. Williams' internalism, I will argue, fails because it cannot account for the normative content of normative claims or beliefs. Normative claims cannot be properly explained in non-normative terms. The term 'reason' is ambiguous and can therefore mislead the theorist who is trying to make sense of the different senses of the term. My main critique

against internalism will be that it conflates normative reasons with motivating reasons. Though I believe that normative reasons are explanatory in one sense, I shall argue that they are not explanatory in the sense that Williams, as we shall see, takes them to be.

What, then, is the relation between our reasons and our motivations? My answer is that there is a normative relation between them. It is true of us that we should want or do certain things, or else we are being irrational. We are, because we can reason, open to rational criticism. To be rational is, I will argue, to be able to recognize and respond to reasons. To have this ability is to be generally rational, and if one is generally rational, it is meaningful to ask, in given situations, whether one is being rational or not. As only persons, as far as we know, have the rational ability to reason, it is only they that can be rational or irrational.

While the rationality of our actions depends on our beliefs, what reasons we have does not. Thus if I have false beliefs, I may be rational in doing what there really is no reason for me to do. The belief in question may be irrational though. I shall discuss four different ways in which beliefs can be rational or irrational. The rationality of our beliefs is not solely dependent on their relations to reasons. Beliefs can also be irrational because they aren't consistent with other beliefs that we have. Thus, it can, for the sake of consistency, be rational to believe what there is no reason to believe.

Both beliefs and desires can, I believe, be intrinsically rational or irrational. It is usually easier to tell when beliefs or desires are intrinsically irrational. They are irrational in this sense when the content of the belief, or the object of desire, is such that it is obvious that we don't have reason to believe, or desire, what is believed, or desired. These claims are claims that can only be made on the kind of view that I am defending.

If it is intrinsically rational to desire something, then the object of desire is of value. Thus we can say that claiming that something is good is basically another way of saying that there is reason to care about it for its own sake; something is good if it has features that give us reasons to care about it.

I shall start out by discussing reasons. Having done that I will, in the light of the conclusions I draw about reasons, turn my attention to rationality. I call the view that I will be defending here *Normative Realism*. Externalists about reasons don't have to take this view. There are, for instance, also constructivist<sup>2</sup> views about external reasons that are promising, but I shall not discuss such views or why I think that we should reject them here. Nor shall I defend normative realism against some of the objections that may be raised against such a view.<sup>3</sup> My way of arguing for normative realism will thus be indirect

in the sense that I present it as, what I take to be, a better alternative to the internalist view defended by Williams that I below will present and reject.

#### *Different kinds of reasons*<sup>4</sup>

Since the word 'reason' is so ambiguous, I first have to say a little about certain distinctions and the different senses of 'reason' that will be important in the discussion. While doing this I shall try to point out where the problems and the disagreements are found. We can begin by considering the following situation:

*The fire:* Unaware of the fact that my hotel is on fire, I am about to go to sleep. If I hadn't been using earplugs in order to shut noise out, I would most likely have heard the people panicking in the hallway. Thus instead of saving myself by jumping into the canal outside my window, I decide to try to sleep, as I am feeling quite tired.<sup>5</sup>

I would like to make the following three remarks about this situation: First and foremost, it appears clear that I here, even though I am unaware of it, have a good reason to jump out of the window. Second, my failure to do what I have reason to do is, since I am not aware of the fact that the building is on fire, not a case of irrationality. And finally, the explanatory reason for my going to bed seems to be that I am tired. The first claim is a claim about what normative reasons I have, the second claim concerns the rationality of my acting and the third claim aims to explain why I act the way I do in this given situation. It is the interpretation of the first claim that I, in this paper, will say most about.

When we speak of normative reasons for acting, desiring or believing we attempt to justify our acts, desires and beliefs. As reflecting beings most of us want to be able to justify how we live, not only to others, but to ourselves as well. Thus we also think about reasons in order to find motivation. If I don't believe that there is any reason to do x, or if I believe that there are reasons against doing x, then it is unlikely, if I am rational, that I will be motivated to do x. Alternatively if I believe that there are reasons to do x, it is likely that I will be motivated to do x, as this belief will make x appear meaningful.<sup>6</sup>

When we make claims about reasons we can, roughly, be said to be trying to answer questions starting with 'why'. Think of questions like "why should I jump?", "why do you believe that the building is on fire?" and "why do you want to jump?". The answers to

these questions are all reasons. Thus reasons, besides justifying, also *explain*. They explain either why we *ought to*, or why we *do*, act, desire or believe various things. This distinction between why, on the one hand, we ought to act, desire and believe, and why we, on the other hand, do act, desire and believe, is a distinction that, throughout the course of this paper, will be of great importance.

Now, how does what I have just said apply to the example given above? Suppose that the phone rings and that you are on the other end. You tell me that I should jump out of the window and, baffled by this, I reply “why?”. You say that I should jump because the building is on fire. I reply by asking why you believe that the building is on fire and you answer by saying something like “because I see flames and smell smoke”—you are in the building across the street we can suppose—and thereby offer two reasons to believe that the building is on fire. Because I take these to be good reasons to believe that this is true and because I take the fact that the building is on fire to be a good reason to jump out the window, I now want to jump. Failing to want to jump upon receiving this information seems, as the fact that the building is on fire appears to be a strong reason for me to jump, to be irrational; jumping would save my life.

Now, the fact that the building I’m in is on fire seems, then, to be good a reason for me to jump. If I stay in my room I will either be badly burned or die. If I am aware of the fact that jumping would save my life then I will, hopefully, be *motivated* to jump out the window. If that is the case, then my motivating reason is the content of my belief that jumping will save me, while the normative reason is the fact that jumping will save me.

Sometimes we do things that we have no good reasons to do. Thus the idea of motivating reasons is distinct from that of normative reasons. This would, as should be kept in mind, be true even if we always did what we have good reasons to do. And, if there were no such things as normative reasons, we could still explain people’s acting.

When we speak of *explanatory* reasons our aim is, as the name implies, to explain actions. Our explanatory reasons are our motivating reasons when we act intentionally. When we act intentionally we can answer questions about why we act as we do. Though we often take our motivating reasons to be good reasons, we do sometimes admit that our reasons for acting aren’t good ones; as rational beings we know that we sometimes act irrationally.

Keep in mind also that we sometimes ‘do’ things unintentionally. At such times our actions should not be explained by appealing to motivating reasons. We ought then only to speak of explanatory reasons. Why, for example, did Tim walk about the house like a

zombie last night, not answering when spoken to? The answer, or explanation, is that he was walking in his sleep. Here it seems awkward to speak of any motivating reasons. Acting on reflexes also excludes motivating reasons and, as it is possible to explain such acting, we can conclude that motivating reasons and explanatory reasons aren't necessarily the same thing<sup>7</sup>.

We are, as persons, able to recognise normative reasons and to respond to them. Our motivating reasons are then the same as our normative reasons in the sense that they both depend on the same fact(s)<sup>8</sup>. When we act rationally and have no false beliefs—the rationality of our actions depends, as I will get back to, on our beliefs—our normative reasons will be our motivating and, thus, also our explanatory reasons<sup>9</sup>. So, even if the different kinds of reasons that we are discussing coincide in a way at times when we are rational, we here have three separable kinds of reasons: Normative reasons are facts we ought to be motivated by. Motivating reasons explain our actions when we act intentionally and are thus explanatory reasons. When we act unintentionally, for example when we act on reflexes, there are still explanatory reasons—like the fact that we have certain reflexes—that explain why we acted as we did.

Explanatory reasons can be relevant to appeal to in causal explanation. A person's being in some mental state may cause her to act in a given way. If we accept this, we hold, what Jonathan Dancy<sup>10</sup> calls, a *psychologist* view. When we explain actions by appealing to explanatory reasons, we take the third-person perspective. When it comes to motivating reasons, it is the content of these mental states—or, more exactly, our beliefs—that is relevant. The first-person perspective is, in explanations in which motivating reasons are appealed to, what is taken. When we are discussing the rationality of people's acting, desiring and believing it is motivating reasons that are most relevant. There is, as should be recognised, no conflict involved here. Consider an example: Because John believes that he will be able to catch his train if he runs to the station, he runs. The explanatory reason for him running was that he was in a particular mental state. His motivating reason was the content of his belief, i.e. that running to the station would make it more likely that he would be on time. Thus, if we ask him why he is running, he will answer that he has to in order to make it to the train on time. He will not answer that he runs because he was in a particular mental state before he started running. This will be important later on.

That we can distinguish between normative and motivating reasons is widely accepted. It is the nature of the relation between these kinds of reasons that creates controversy.

### *Good reasons and strong reasons*

The following is what I, when I say that there is reason to do or want something, will be meaning: To say that something is a normative reason speaking in favour of doing or wanting something, is to say that doing or wanting this thing for this reason would be to do or want it for a *good* reason. Thus whenever we say that there is reason to do or want something, we don't need to say that there is a *good* reason<sup>11</sup> to do or want this thing; that is implied when we say that there is reason to do or want it. If we say that something is a good reason, we don't add anything.<sup>12</sup>

When we say that someone's reason for doing or wanting something is a bad reason, we are basically saying that there is *no* reason for doing or wanting this thing. The person may, if she has got false beliefs about the thing in question, not, we could grant, be irrational in doing or wanting this thing. So, to say that there is no reason for someone to do, or want, something is not necessarily the same as saying that anyone who does, or wants, this thing is being irrational. Thus when Louie the fourteenth had all of his teeth removed because he (for some obscure reason) thought that it would be good for him, he wasn't, even though there was reason to abstain from doing this, necessarily being irrational. He may have had some false belief about teeth whose truth could have given him reason to want to do this. Besides, there may actually have been some reason speaking in favour of pulling out all of your teeth in those days. Maybe it showed that you were rich enough not to have to chew. But if this actually was a reason for getting rid of one's teeth, it could not have overridden the reasons against doing this. When there are both reasons speaking against and in favour of doing, or wanting something, we can ask what reasons are the *strongest*. Something can be a reason without being a strong reason; if something *is* a reason to do or want something, it is still, even if it is a weak one, what we, on my understanding of these concepts, can call a good reason to do, or want, this thing.

### *The internalism/externalism-distinction*

I have, up until now, tried to leave it open how we, more exactly, ought to understand claims about normative reasons and the nature of such reasons. We shall now, in order to get our discussion about this really started, consider Williams' very influential view on the relation between normative reasons and motivation. Williams makes a distinction between internalism and externalism about normative reasons. On internalism, which is the view

that he favours, there is a necessary connection between an agent's reasons and her motivations. Contrary to Williams I believe that we should accept externalism, which, simply put, is the denial of this supposedly necessary connection. When Williams first introduced the distinction in question, he made the following presentation:

Sentences of the forms 'A has a reason to [do x]' or 'there is a reason for A to [do x]' (...) seem on the face of it to have two different sorts of interpretation. On the first, the truth of the sentence implies, very roughly, that A has some motive which will be served or furthered by him [doing x] (...). On the second interpretation, there is no such condition, and the reason-sentence will not be falsified by the absence of an appropriate motive.<sup>13</sup>

We shall, following Williams, call followers of the first interpretation *internalists* and followers of the second interpretation *externalists*. We can also, at first glance, note that this passage seems to imply that what we are discussing here are solely questions about meaning. Thus, in a later article on the same subject matter Williams asks: "what are we *saying* when we say someone has a reason to do something?"<sup>14</sup> and later, in the same article, he says that he thinks that "the *sense* of a statement of the form 'A has reason to [do x]' is given by the internalist model"<sup>15</sup>. But, when we ask, as Williams also does, what the truth conditions for reason statements are, and when we ask what claims about reasons are *about*, these questions cannot, I believe, merely be questions about language. We also want to know what a reason *is*, and in what sense reasons, if they 'exist' in any sense, exist. Are there, in other words, any reasons?

### *Normative Realism*

Beliefs about reasons, we should be aware, could be illusions; there may not be any such thing as normative reasons. This, I believe, is a genuine possibility. On the internalist view, it may be objected, this possibility is excluded. What this view claims is rather that some of our beliefs or claims about reasons are false or make little or no sense. A true reason statement or belief is not a statement or belief about some external fact or truth; a true statement or belief to the effect that somebody has a reason implies, rather, that a certain action is one that an agent could, or would, be motivated to perform under certain conditions. Thus it might be held that we, on internalism, don't have beliefs 'about'

reasons, but beliefs about the relation between certain actions and their relations to our motives/motivation. And, when no such relation holds, it is correct to say that the agent does not have reason to perform this action.

We should, I believe, reject such claims; claims and/or beliefs about reasons are, or so I shall argue, not claims about relations between actions and motivations in the sense just described. On the externalist view that I believe we should accept it is true that beliefs can be irreducibly normative—we assign the status of ‘reason-hood’ to certain facts—and that there are irreducibly normative truths<sup>16</sup>—some of these beliefs are true. Accepting internalism would, on this view, be to believe that there are no normative reasons. As I believe that there are normative reasons in the non-reductionist sense just mentioned, I call the view that I am defending ‘Normative Realism’<sup>17</sup>.

Since the internalist view may seem less controversial, I must, in order to make a case for normative realism, try to show why we should not accept internalism. Let us, therefore, get to that.

#### *The first formulation of the internalist view*

Formulated as in the Williams quote above, the internalist view may appear unclear. The inclusion of the term ‘motive’ is what I, especially, have in mind. Motives, taken at face value, appear to be what I have called motivating reasons. Motivating reasons, I have said, are contents of our beliefs that motivate us. If motives were the same as motivating reasons they would, then, be beliefs. Williams, however, takes motives to be members of, what he calls, an agent’s motivational set (or the agent’s S), which may include the agents “desires, evaluations, attitudes, projects and so on”<sup>18</sup>. With the possible expectation of evaluations, few of these so-called motives appear to be beliefs. What we have here are rather desires, pro-attitudes and/or intentions. So ‘motives’ are, then, not the same thing as that which I have called motivating reasons. One cannot ‘serve’ or ‘further’ one’s motivating reasons, they are simply things that motivate us. For this reason, I shall take motives to be motivations or desires.

To explain the internalist view Williams assembles four propositions that he takes to be true of internalism<sup>19</sup>. What I have said above about internalism implies the first one, which reads:



- (i) An internal reason statement is falsified by the absence of some appropriate element from S

Given the formulation above it appears to be the case that we, according to the internalist, have reason to serve or further all or any of our motives/motivations. That this is deeply implausible is recognised by Williams. He considers an example in which an agent desires a gin and tonic<sup>20</sup>. The agent thinks that the stuff in front of him is gin and wants, therefore, to mix it with tonic and drink it. However, the stuff is in fact petrol. As Williams notes, it appears clear that the agent does not have reason to act on this desire. The reason Williams offers for this is the following: had the agent known that it was petrol and not gin that he has in front of him, it is most likely that he no longer would desire to mix it with tonic and drink it. From this Williams extracts:

- (ii) A member of S, D, will not give A a reason to [do x] if either the existence of D is dependent on false belief, or A's belief in the relevance of [doing x] to the satisfaction of D is false.

From this Williams deduces the following epistemic consequence:

- (iii) (a) A may falsely believe an internal reason statement about himself, and (we can add) (b) A may not know some true internal reason statement about himself.

(b) is partly explained by Williams' fourth proposition which reads:

- (iv) internal reason statements can be discovered in deliberative reasoning

Williams' wording here, and in (iii) as well, seems to suggest that the view he has in mind is analytically reductive. Upon having deliberated the agent, Williams claims, may be correct in *saying* about herself that she has a reason to do a certain thing. By this the agent will, it seems, mean nothing more than that doing this thing will serve or further a motive that she has in her S. This would, rather than being a normative claim, be a non-normative claim about the relation between a particular action and the agent's S; it would, I believe, be a claim about a motivating and not a normative reason. If this is all that the correct

application of the word 'reason' amounts to, there would, I believe, be no truly *normative* reasons; there would only be relations between acts and motivations that could properly be described by using the word in question.

Perhaps Williams should have written:

(iv\*) internal reasons can be discovered in deliberative reasoning

This could mean that an agent could, after having deliberated, come to conclude that he has a reason to perform a certain action. Assuming that he accepts internalism, his reasoning could go like this:

(1) I am now, after having deliberated, motivated in a certain way (and, so far as I know, this motive does not rest on any false beliefs)

(2) Doing x would serve or further my motive

(3) We have reasons to do things if they would serve or further our motives, given that our motives don't rest on false beliefs

Therefore (so far as I know):

(4) I have reason to do x<sup>21</sup>

This agent doesn't merely want to know whether it is correct to say about her that she has a reason to act. She doesn't, in other words, want to know whether it could, as another way of describing the relation between some act and some motive of hers, be correct to say that she has a reason. What she wants to know is not how she should use certain words, but rather how she should act; she wants to know whether, or not, she has a reason to act. That she has a motivating reason, or is motivated in a certain way, is something that she already knows. The question is whether she has a *normative* reason to act as she is motivated to act. The claim that she would, after deliberation, be motivated to do x, is not a normative claim. Nor is the claim that somebody believes of herself that she has a normative reason normative.

In virtue of what, we may now ask, is it true that the agent in question has a normative reason to act? We can, in order to try to answer this question, consider two different kinds of internalisms.

## *Internalisms*

Williams did, as we have seen, make the claim that ‘the truth of the sentence [that A has a reason to do x] implies, very roughly, that A has some motive which will be *served* or *furthered* by him [doing x]’.<sup>22</sup> This seems to allow us to, within internalism, make a further distinction, namely between:

*The desire fulfilment view:* A has a reason to do x if doing x would help fulfil desires that A wouldn’t lose after going through deliberation

and

*The deliberative view:* A has a reason to do x if, after going through deliberation, x is what A would be most motivated to do

The first view I get from ‘served’ and the second I get from ‘furthered’. In order to evaluate these views, we can consider less sophisticated versions of them:

*The simple desire fulfilment view:* A has a reason to do x if doing x would help fulfil A’s present desires

*The simple deliberative view:* A has a reason to do x if x is what A is presently most motivated to do

Even though it may appear clear that both these views are implausible, there is one feature that the first of these two views has that makes it a little more plausible than the other: That doing a certain thing would help fulfil a desire of yours could, we should agree, be a reason for you to do this thing. Acting just because you want to appears to be to act for no reason. Sometimes when we are asked why we are doing some particular thing, we simply answer: “No reason—I just felt like it”. On the deliberation view, be it the simple or the sophisticated version, that is a reason. We may ask a person who gives an answer of the kind just mentioned why she felt like doing whatever it is she is doing. Here we may expect a more informative answer; we don’t just come to desire things. Besides, if we did, that doesn’t seem to be something that could be the source of reasons.

Remember now my distinction between explanatory and motivating reasons. There are, even if there aren’t always motivating reasons, always explanatory reasons why we have

the desires we have. I desire to eat because I haven't eaten for a couple of hours. I am tired because I have been up too long. And so on. Sometimes we desire things as a result of our beliefs. I desire to go on a trip to New York City because I believe that I would enjoy it. In such cases we know why we desire what we desire. However, it may also be the case that we don't really know why we desire some of the things we desire. In such cases it seems unlikely that these desires could provide reasons for us to act.

Reasons, I have claimed, do two jobs: They *justify* our acting, desiring and believing, and they *explain* why our acts, desires and beliefs can be, or are, justified. Can internal reasons justify and explain why we ought to care about certain things or do certain things? Consider first the question of justification applied to the deliberative view. After going through sound deliberation, what I most want to do is x. Does this justify my doing x? It may be the case that I am justified in doing x, but does my wanting to do x under certain conditions constitute this justification? Somehow it seems reasonable to ask *why* I am motivated to do x. What considerations made me want to do x? This, I believe, has to be what is relevant. If I, for instance, want to do x because this would prevent me from great suffering, it appears clear that I can justify my wanting to do x. That doing x would prevent me from suffering can also explain why this motivation of mine is justified. This, to me, strongly suggests that the deliberative view is mistaken. What's important, or relevant, is *why* I am motivated in certain ways.

Consider now the desire fulfilment view. That I have a reason to do x is, on this view, true if doing x would help fulfil the desires I would have after having deliberated soundly. This has more explanatory value. That doing x would help fulfil a desire of mine explains some of the point of doing x. But, if it is to justify my doing x, it is still important to know whether the desire that doing x would help fulfil is one that I can justify and here we can again appeal to the points I just made about the deliberative view. We might also ask why, or whether, it is a good thing that some desire is fulfilled. Insofar as fulfilling a certain desire would bring me some kind of satisfaction it certainly appears true that I could have reason to fulfil this desire. But the reason for me to do x then appears to be that doing x will bring me pleasure, rather than that doing x would fulfil my desire. And, it cannot, moreover, be true that it is always good to fulfil our desires. Some of us desire to hurt other people. These people may not lose these desires after finding out what is relevant to know about other people's pain. On the desire fulfilment view these people would have reason to start hurting other people. That cannot be true. Surely they will have motivating reasons, but such reasons should not, as I have argued, be conflated with normative reasons.

Williams does, as I've already said above, reject the simple views. This implies that he takes the element of deliberation to be of great normative importance. It may, therefore, be there that we find the element of normativity that we have been looking for.

*A better formulation of the internalist view*

In a postscript to *Internal and External Reasons* Williams writes:

The formulation of the internalist position that I now favour is: A has a reason to [do x] only if there is *a sound deliberative route* from A's subjective motivational set (...) to A's [doing x].<sup>23</sup>

I agree with Williams that this is a better formulation. In the first formulation considered above normative and motivating reasons are conflated in a way which appears to do away with normativity. Williams may have been moved by this kind of consideration when he got convinced that there was a better way of expressing the internalist thesis. He writes:

It is important that even on the internalist view a statement of the form 'A has a reason to [do x]' has *normative force*. (...)A has a reason to [do x]' means more than 'A is presently disposed to [do x]'. One reason why it must do so is that it plays an important part in discussions about what people should become disposed to do.<sup>24</sup>

This may be taken as being trivial; that we cannot have normative discussions without normative concepts is undeniably true. Williams may have had something more important in mind. Consider:

*The appeal to the need for normative concepts:* It may be true that we actually, in a non-trivial sense, have reasons to want, and do, certain things. If we don't have any concepts that are truly normative, we cannot discuss questions about these things. Neither does it seem that we could have any normative beliefs. Therefore, we need concepts that are truly normative.

This appeal to the need for normative concepts is not trivial. The question now is whether the internalist view really is normative in the relevant sense.

Consider these two claims that could be made by internalists:

C1: A has a reason to do x if A *could*, after going through sound deliberation, be motivated to do x

C2: A has a reason to do x if A *would*, after going through sound deliberation, be motivated to do x

While C1 speaks of a possibility to reach the conclusion to act, C2 speaks of a necessity. We can start by considering C1.

C1, it seems, leaves it, as 'could' allows for an extremely wide range of possibilities, rather open what the agent has reasons to do. Thus this view appears, we might say, to be unacceptably relativistic. This does, of course, depend on what counts as 'sound' deliberation. As we shall see, this, i.e. that the view is rather relativistic, will be true if we accept Williams' view of what sound deliberation is. It also, as we can note, seems to be the case that all it takes for it to be true that an agent no longer has reason to act in a given way is a change of heart. As this is true of C1, this formulation of internalism may seem hopelessly vague and open.

When it comes to C2 it will not, it seems, be left as open what the agent has reason to do. This, again, has to do with what counts as sound deliberation. It also has to do with what psychological assumptions one wishes to make. If we, for example, accept psychological hedonism, i.e. the belief that we ultimately always only seek pleasure, then it is only our beliefs about what best achieves this aim that will put constraints on, or will set limits to, the range of possible reasons we may have as a result of sound deliberation. However, taking into account the so-called paradox of hedonism, which is the very probable truth that we will be less successful in achieving our hedonistic aim if we only seek happiness (and nothing else) for its own sake, adherents of this view may claim that we have reason to seek other things for their own sake. Thus it will, given that different things bring different people pleasure, be true, just as internalists claim, that our reasons are relative to our subjective motivational sets. This may, however, to some extent also be true on externalist theories. It may, for example, be true that I have reason to do or want something that you have no reason to do or want, because, given our preferences, I will get

pleasure from doing this particular thing while you will not. The difference is that on externalism we have these reasons because of these latter facts about our possible pleasure or pains. And, these reasons are grounded in the desirability of pleasure and/or the (normative) fact that pain is something that is worth avoiding.

If the internalist accepts C2, then, we might ask, why should he reject externalism? ‘Sound deliberation’ appears, on the face of it, to involve at least prudential considerations and, as moral rationalists claim, moral considerations as well. What we need to know, then, is what Williams means when he says that the deliberation ought to be ‘sound’. On the reading just suggested the distinction between internalism and externalism appears to come undone. The externalist could, as Parfit points out<sup>25</sup>, claim that it is true that, if we have reason to do x, *substantively* rational agents *would* be motivated to do x after deliberating soundly. To be substantively rational, it isn’t enough to be able to deliberate rationally, one must be motivated in certain rationally required ways, and, if presented with the opportunity, actually act on these desires.

There are, then, two questions that we first need to know the answers to, namely:

Q3: Should the internalist accept C1 or C2?

Q4: How should the internalist understand the idea of a ‘sound deliberative route’?

Williams comments on this in the following way:

[W]e cannot necessarily equate ‘A could arrive by sound deliberative route at the decision to x, with ‘A would arrive at that conclusion if he deliberated and did so soundly’. Someone may, indeed, do what he has most reason to do without deliberating and, perhaps, because he did not deliberate. We must not assume that thinking about the question about what to do, rehearsing considerations, and so forth, are simply like the perception of an external reality.<sup>26</sup>

Thus it appears that the internalist has to embrace C1 rather than C2. But that is, I believe, problematic. It is a problematically contingent question whether an agent would reach a certain conclusion after going through sound deliberation. This, of course, has to do with what one takes ‘sound’ to imply. The risk here, for the internalist, is that we accept a too

broad interpretation, one which includes substantive constraints like the recognition of certain values or prudential or moral considerations. That, as I have said, would seem to undo the distinction between internalism and externalism. Externalists takes substantively rational agents to be agents that recognise certain aims as being worth achieving, or, alternatively worth avoiding and/or preventing. Even though he may not count as externalist we can note that Kant did, when he considered what it is to be a rational agent, claim that such an agent necessarily is committed to seeing himself acting impartially. On Williams view, this is far from clear. He remarks:

[The internalist account] demands that the conclusion should be argued, and it cannot be acquired on the cheap, for instance by pointing out the obvious truth that people often describe unfair conduct as ‘unreasonable’. It takes more than ordinary language to deliver large Platonic, Aristotelian, or Kantian conclusions to the effect that that virtue and reason will coincide.<sup>27</sup>

But, as must also be noted, Williams does grant that a desire to correct factual errors in one’s reasoning is a necessary condition for being a sound deliberator. Concerning this, he makes following remarks:

It may be asked why the agents’ deliberative route can, on the internalist account, be shown to be unsound by reference to factual mistakes, while claims that what the agent is doing is immoral or imprudent do not necessarily count as showing that the deliberative route is unsound. The answer to this is, crudely, that an agent is committed in general to acting in the light of sound information, simply by being a rational agent; included in the S of every rational agent is a desire not to fail through error.<sup>28</sup>

Thus, one cannot, Williams claims, be considered a rational agent if one doesn’t desire not to fail through error. Supposing that this may be true, why shouldn’t we allow the recognition of certain values, moral and/or non-moral, to be necessary in order to count as a rational agent? In making this objection it can be left open exactly what these values are here; we don’t have to say anything substantive about the desires of rational agents. We need only claim that rational agents are agents who care about certain things that are worth caring about. Does it not seem possible that certain normative beliefs arrived at after going



through a sound deliberative route, interpreted as being one without errors of facts, could be irrational simply in virtue of their content? If that is so, then these beliefs could be considered as being intrinsically irrational<sup>29</sup>. Consider this example: I consider how pain feels and reach the normative conclusion that I ought to seek it. If I know how pain feels, my deliberative route is in Williams' sense sound, but it doesn't appear very 'sound' to reach this normative conclusion, does it? Somehow it seems hard to think that it is true that I have reason to seek pain. Some may be tempted to say that no believable agent *would* arrive at this conclusion, but we should not, I believe, assume this. We should, rather, admit that it may be possible that the agent *could* arrive at this strange conclusion, and note that it appears strongly counter intuitive to call the reasoning of this agent 'sound'.

The point, then, is this: if we accept C1, we must accept that the most absurd conclusions about what reasons people have may be true. Normative beliefs cannot be intrinsically irrational, or irrational in virtue of their content. Somehow this seems hard to believe. If we can allow Williams' deliberative constraints on rational agents, then, I ask, what is the further reason for not including other constraints?

### *Williams' examples*

In the course of his discussions of these matters, Williams makes use of a couple of examples that we, in the light of what have said above, now will consider. I have already mentioned the example with the man who wants gin and tonic. In that example internalism, we noted, seemed to get things right. It was obvious that the man in the example did not have reason to mix the stuff in front of him with tonic and drink it. The internalist explanation of this agent not having this reason was that he would have lost his desire to mix this drink if he were to find out that it was petrol and not gin that he had in front of him. I, on the other hand, have claimed that what is relevant to know is, rather whether we under certain circumstances would have certain desires, *why* we have the desires that we have, or that we under other circumstances would have. Do our motivating reasons for having these desires justify our having them? If they do, I claimed, they do so because they explain why we are right, or rational, in being motivated accordingly.

Now, Williams discusses other examples, in which internalism does not appear to get things right. In one example we imagine a man who needs to take a certain medicine. He knows this, but doesn't care. Persuasively he denies that he has any interest in preserving his health. On internalism it is therefore true that this man has no reason to take his

medicine.<sup>30</sup> This, to me, appears strongly counter-intuitive. That he needs the medicine appears to be a reason in itself for taking it; if we ask why this man should take the medicine and it is answered that he should take it because he needs it to preserve his health, then this appears to be a satisfactory answer. It may be the case that he wishes not to take it because he knows that, as he has a terminal disease which causes him to be in constant pain, all that the medicine will do is to prolong his suffering. If this is the case, it might be true that he has reason not to take the medicine. But this reason, again, would not be provided by the fact that he has this motivation, but by the fact that the few days that he could go on living would only involve constant pain. These further details are, however, ones that I have added only in order to show that it need not always, on externalism, be the case that we have reason to take medication. In Williams' example no such details are added and it may be the case that this man could, if he takes his medicine, go on to live a long and rewarding life. If that is the case, then it appears strange to say that he, just because he doesn't care, doesn't have any reason to take the medicine.

We can here also return to the example we began by considering, in which I claimed that I, as my hotel was on fire, had a reason to jump out of the window. With regards to this example it may be objected that it is unsatisfactory because it is under-described. What, it might be asked, if I want to die? This objection, I believe, begs the question in favour of internalism, as it seems to presuppose that what I want determines what reasons I have. The point of the example was to illustrate a situation in which it appeared clear that I, even though I was unaware of it, had a reason to do something. If the fact that the building was on fire came to my attention, it is most likely that I would be motivated to jump. But that would be because I strongly would believe that I did have a reason to do so. If it was the case that I wanted to die, we could again ask whether this desire was rational. It may even be the case that, even though my supposed desire to die might be rational, I have reason to jump. Being burned alive does not appear to be a very pleasant, or dignified, way of dying. Just as the man in the last example could be rational in wanting to stop living because he suffered tremendously from disease, this may be what motivates my desire to die. And, I may want to die in a more peaceful way by means of euthanasia aided by doctors who see to it that I die in a painless and planned way. Such a desire could, I believe, be rational.

In another example, which I take to be the most counter-intuitive one, Williams imagines a man who treats his wife badly. Williams tells this man that he has a reason to treat his wife better. The first reason he comes up with is: 'Because she is your wife'. This

has little effect on the man who replies by saying that he doesn't care about that. Williams is persistent and tries to give other answers, but he finds nothing in the man's motivational set that gives him an internal reason to start treating his wife better. Many things may, as Williams concedes, be said about this man. We may call him ungrateful, inconsiderate, hard, sexist, nasty, brutal, etc. But, Williams claims, we cannot say that he has a reason to change. This line of thought seems, we can note once again, to be dependent on a conflation of normative and motivating reasons. It might, indeed, be agreed that this man has no motivating reason, or motivation, to do what we want him to do, but it seems obvious that he has a normative reason to treat his wife better: she suffers from his treatment. And, that we call him all those things that Williams mentions has, it seems, to do with our sensibly taking there to be a reason for him to change his ways; his failure to recognise this reason is exactly what makes him ungrateful, inconsiderate, hard and so on.

That the externalist will claim that this man has this reason is recognised by Williams. But, despite this, he claims, concerning the expected reason-statement, that "one of the mysterious things about the denial of internalism lies precisely in the fact that it leaves it quite obscure when this form of words is thought to be appropriate"<sup>31</sup>. If Williams really, as he claims, thinks that externalism leaves it obscure when it is appropriate to say that somebody has a reason, then how does he know that the externalist will say that this man has this reason? He should, if he thought that the idea of external reasons really is mysterious, have said that he has no idea of what the externalist would say about this man and his reasons. We can also note that Williams, in claiming this, again seems only to be concerned with how the word 'reason' is to be used properly. This is strange. When we consider this case what we think is not merely that it appears appropriate, as a way of describing this situation, to use 'this form of words'. What we consider, or have beliefs about, is, rather, the following normative question: Does this man have a reason to change his ways? Most of us would say that it obvious that he does. Our beliefs, then, about this situation are, in other words, not merely beliefs about language; they are *normative* beliefs, which is a distinct kind of beliefs.

Williams then asks:

[I]f it is thought to be appropriate [to say that this man has this reason], what is supposed to make it appropriate (...)? The question is: 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.<sup>32</sup>

As I have already said, it is, I believe, the very fact that we believe that the man in example fails to respond to the reason he has that makes us say those other things about him. This is the difference between saying that he has a reason and saying those other things about him. And, to say that it is better to do what one has reason to do is, or so I take it, to trivially say that one has reason to do what one has reason to do. The question whether we have reason to do what we have reason to do is just plainly silly. What makes it appropriate to say, or to believe, that this man has reason to change his ways is the normative fact that the well-being of others matters. This, I believe, is best taken as being an irreducibly normative truth. We need not give any further explanations why the well-being of others is something that we should care about. Nor is any such explanation possible to give. Yet, it is a fact that the well-being of people matters. These two claims, namely

(M) The well-being of others matters

and

(R) We have reason to care about the well-being of others

do not seem to report different things. Rather, it seems that to say that the well-being of others matters, is to say that we have reason to care about the well-being of others. If we believe (M) and (R), these beliefs must, I have claimed, if they are to be really normative, be irreducibly normative, and if these beliefs or claims are true, then it is appropriate to say that this man has reason to start treating his wife better.

So, the reason why the man in the example should change his ways is the fact that the way that he treats his wife causes her to suffer. But, the fact that his behaviour has this effect on her is not a normative fact, but a non-normative empirical fact. This non-normative fact is, however, one that has normative significance. Compare:

(E) The way this man treats his wife makes her suffer

with

(N) Because (E) is true, this man has a reason to stop treating his wife as he

does

Although (E) is the fact which is the reason why this man should change, (E) is not a normative claim. Understanding and knowing (E) to be true does not, therefore, necessarily involve understanding and knowing (N) to be true. But, understanding and knowing (N) to be true would involve understanding (E).

There may be different kinds of reasons for caring about the well-being of others. Many would accept that we have prudential reasons to care about others, since if we do, then our own lives are likely to go better. We could, in other words, have purely egoistic reasons to care about others. It may, though, seem more obvious that we have moral, or altruistic, reasons to care about others. We should, in other words, care about others for their sake and not (merely) for our own. There need not be, we should note though, any conflict involved here. We can have reasons to care about others for both their and our own sake. That we, as the internalist wants us believe, only have reason to care about the well-being of others if we *do* care about them or are motivated to promote their well-being, seems hard to believe.

### *Normativity*

If we suppose that we can be able to explain, using only non-normative concepts, what it is to say that the man in the example above has reason to treat his wife better, we are likely, like Williams, to find claims about external reasons mysterious. But, if we do not wish to abandon the normativity of normative concepts and/or claims, we should give up this hope (if we have it). We can appeal to:

*The normativity of normative claims:* Because normative claims are *normative* they cannot be reduced to *non-normative* claims

This claim might seem trivial, but it shouldn't be taken that way. Rather, it should be taken as saying something important about normative concepts. I shall now explain why this is so.

Now, Williams takes the if-part of claim C1 above to have normative force. This was the claim that

C1: A has a reason to do x *if A could*, after going through sound deliberation, be motivated to do x

Is the 'if' in C1 relevantly normative? As already indicated, and as Williams notes, all the normativity to be found in a claim like C1 must, on the internalist view, be found in the condition of there being a sound deliberative route from the agent's present motivations. As I remarked above, this may, given Williams' idea of sound deliberation, seem like a surprisingly weak and narrow view of what normativity consists in, or where, to borrow an expression from Korsgaard, the source/sources of normativity is/are to be found. Williams writes:

Unless a claim to the effect that an agent has a reason to [do x] can go beyond what the agent is already motivated to do – that is go beyond his already being motivated to [do x] – then certainly the term will have too narrow a definition. 'A has a reason to [do x]' means more than 'A is presently disposed to [do x]'.<sup>33</sup>

As Williams points out, it is obviously false that a claim like 'A has a reason to do x' merely states that A presently is disposed to do x for that is only to say that the agent has some motivating reason do x. So, if we were to say that 'A has a reason to do x' means that x could, after going through sound deliberation, still be motivated to do x, would we include the important element of normativity that Williams doesn't wish to lose? In order to try to answer this question we can consider the following remarks of Williams' on the issue (which, in part, repeats a point that we've already discussed):

The claim that somebody *can* get to the conclusion that he should [do x] (...) by a sound deliberate route involves, in my view, at least correcting any errors of fact and reasoning on the agent's view of the matter. (...) We are allowed to change- that is improve and correct- his beliefs of fact and his reasoning in saying what he has reason to do. That is already enough for the notion to be normative.<sup>34</sup>

As all we are saying is that A *could* be motivated to do x after having deliberated soundly, it does not seem to be the case that we are saying something relevantly normative, or something that has much normative force. Williams' view on what normativity consists in appears to be far too narrow. Differently put, we can say that we seem, on internalism, to

have done away with the distinctively normative force of the claim that A has a reason to do x.

### *Naturalism*

If we take ‘A has a reason to do x’ to *mean* ‘when knowing all the relevant facts x is what A is most motivated to do’ or ‘doing x would help fulfil a desire that A wouldn’t lose after knowing all the relevant facts’, then our theory will be analytically reductive. Such a view, we should agree, has little appeal. Therefore we might instead say that, though these sentences don’t have the same meaning, they report the same fact. This would then be a non-analytical but still reductive version of internalism. Both analytical and non-analytical versions of internalism are naturalist views.

We should, I believe, reject all forms of naturalism. What distinguishes normative claims, and concepts, from other claims, and concepts, is, as I have claimed, precisely and notably the fact that they are *normative*, whereas non-normative claims or concepts are *not*. We should, therefore, reject naturalism. Though this is far from a knock-down argument against naturalism, I do, as I’ve said, believe that there is reason for thinking that this claim isn’t trivial. What I think that this claim does is to summarise, or catch the essence of, three other arguments against naturalism.

First we can, using Moore’s strategy<sup>35</sup>, say that it appears to make sense, upon hearing that somebody is motivated in a certain way after going through a sound deliberative process, to ask whether he really has a reason to act as he now is motivated to act. It is, we could claim, an open question whether this person has this reason; it could, even by those who fully understand the concepts involved in making the claim that Williams makes, be intelligibly questioned whether the person has this reason. This, we might claim, shows that the internalist formula for what it is to have a reason isn’t a definitional truth, and thus, we could claim, the analytically reductive version of internalism is refuted.

The internalist may reply that we should revise our use of ‘reason’ so that it no longer becomes an open question whether someone who, after having deliberated soundly on the facts, is motivated to do x, has reason to do x. Following Williams it might be claimed that this is the only intelligible sense in which the term ‘reason’ can be used. But, this would not be a good reply. It assumes that we need a reductionist view and that we cannot understand what ‘reason’ means if it is taken to be an irreducibly normative concept. We

need not, as I've argued, assume this. Irreducibly normative concepts can and, we should agree, *do* make sense.

Further we could, inspired by Sidgwick<sup>36</sup>, say that if we, by the words 'A has a reason to do x', mean no more than that 'A would, under certain circumstances, desire to do x', we cannot be claiming to say anything significant when we say that the internalist view is true. If these two claims meant the same thing this would mean that I, if I had accepted the internalist view, could not in any interesting sense reason in the following way:

- (1) I am now, after having deliberated on the facts, motivated to do x.
- (2) I, therefore, have a reason to do x.

(2) merely would be a restatement of the premise (1). My reasoning would be unacceptably trivial. This argument hits both analytically and non-analytically reductive views. On a non-analytical view (2) does, even though it doesn't have the same meaning than (1), report the same fact as (1) and can therefore not be taken to be a significant conclusion. Parfit, when discussing this<sup>37</sup>, concludes that reductionists are, when trying to explain normative reasoning, committed to the following kind of reasoning:

- (3) Jumping would get me what I most want.
- (4) As another way of reporting this fact, I could say that I have most reason to jump, or that I should jump.

Since (4), even though it by these internalists is taken to be a normative claim, does not report any fact that is different from (3), they must accept that (4) is the only thing they can infer from (3). In a discussion of this argument, Parfit makes the following remarks:

Like the analytically reductive view, this account grotesquely distorts my reasoning. When I conclude that I should jump, I am not merely redescribing my argument's premise. I am drawing a quite different conclusion. If my conclusion merely redescribed the causal fact described in (3), it could not be normative.<sup>38</sup>



The consequence of accepting naturalism, Parfit claims, is that we abandon normativity. Naturalism fails, in other words, because it describes and/or explains our normative beliefs in a way that abolishes the normative content of these beliefs.

It may now be objected that while this may show that normative concepts aren't reducible, it does *not* show that normative properties cannot be natural properties. They could be some sort of natural properties that aren't reducible to natural properties because they supervene on natural properties. Because something has certain natural and non-normative properties it has some normative natural property. This could, as it appeals to the irreducibility of normative properties, be considered to be a realist view. Among the properties of things there are normative natural properties that are just as much part of the fabric of the world as any other properties. We gain knowledge of these properties by causal interaction with them, just as we gain knowledge of other real properties by means of causal interaction with them.

Such a view, I believe, fails because it misunderstands the nature of normative properties. There is a deep difference between natural and normative properties that we should recognise. Our normative beliefs do, as well, differ from our non-normative beliefs in a significant way. What I mean by this will become apparent as we go along. I will not argue against naturalist realism any further here<sup>39</sup>. For now I will assume that this kind of naturalist realism fails, and get back to our main target, which is internalism.

### *Non-reductive internalism*

Shouldn't we, then, claim that on a truly normative version of internalism, 'A has a reason to do x' has the implication that if A, after having gone through sound deliberation, would be motivated to do x, then a further, and distinct, fact would obtain, namely the normative fact that A *should* do x? It would then be the case that the counterfactual claim about A's possible motivation(s) would have normative force because it would be the source of normativity. This would be a non-reductive form of the internalist view, one on which the crucial element of normativity is not vanished. Because certain motivational facts obtain, another kind of fact obtains as well: we have reasons to act in certain ways. Parfit, who considers this possibility, suggests the following formulation of the view at hand:

(A) Some acts really are rational. There are facts about these acts, and their relation to our motivation, which give us reasons to act in these ways.

He compares this with realism, on which the following is true:

(B) Some aims are really worth achieving. There are facts about these aims which give us reasons to want to achieve them.

A lot of writers have, Parfit notes, been sceptical towards claims like (B) because of their metaphysical implications; (B) commits its defender to believing in irreducibly normative truths. The same is, however, true of (A). (A), being non-reductive, claims that since certain motivational facts obtain, other normative facts obtain as well. By making this claim, the non-reductive internalist commits herself to at least one kind of irreducibly normative truth (namely the one just mentioned). This undermines her reason for claiming that there can be no truths about what is worth achieving. Thus, Parfit concludes, (B) is no less plausible than (A).<sup>40</sup>

The following is meant to be an alternative version of the argument just presented: Suppose the internalist appeals to an ‘ideal-observer’ version of internalism. Then she might say the following:

(C) Some acts really are rational. These are the acts that an ideal observer would be motivated to perform, or would approve of, if he knew all the relevant facts about given situations.<sup>41</sup>

These acts might be acts that Hare’s archangel<sup>42</sup> would be motivated to perform. (C) incorporates an element of motivation by stating that we have reason to do what the ideal observer would want us to do. The argument could here be that, while it is true that (C) speaks of a motivational element, this addition is redundant. What is relevant is, as I have already argued, *why* the ideal observer would be so motivated. It seems that it would be enough just to say that we have reasons to act in certain ways, without adding that the archangel would be motivated by certain beliefs or facts. And as (C), like (B), commits its defender to belief in at least one normative truth (i.e. the truth of (C)), (B) would be no less plausible than (C).

The archangel, we can suppose, is *substantively* rational. Williams appeals, as we have seen, only to *procedural* rationality. Claiming that somebody’s reasoning is not a ‘sound deliberation’ is at least mildly normative, as the appeal to ‘soundness’ is slightly

normative<sup>43</sup>. But, since what, on Williams view, seems to be of most importance is the process of deliberation that leads somebody to think that he should act in a certain way, his theory is perhaps best seen as a theory of practical *reasoning* rather than of practical *reasons*. Williams, I believe, would reject both (A) and (C). He would claim that, though there are no acts that we have any actual reasons (in the sense that (A) and (C) seems to imply that there are) to perform, there are ways of coming to believe that we have reasons to act in certain ways that are less rational or desirable than others. Williams himself writes that “the internal reasons conception is concerned with the agent’s rationality”<sup>44</sup>. In some ways, then, his theory seems to be only about instrumental rationality.

### *The relation between reasons and motivation*

I have claimed that internalism conflates normative and motivating reasons. Williams would, of course, disagree. In fact, one of Williams’ arguments in favour of internalism does focus on the relation between these two kinds of reasons. We shall now examine that argument. After first having quoted what Williams says in its entirety, I will, step by step, go through that which he says, pointing out several things that I take there to be reason to reject. Williams writes:

The (...) point is the interrelation of explanatory and normative reasons. It must be a mistake to simply separate explanatory and normative reasons. If it is true that A has a reason to [do x], then it must be possible that he should [do x] for that reason; and if he does act for that reason, then that reason will be the explanation of his acting. So the claim that he has a reason to [do x] (...) introduces the possibility of that reason being an explanation; namely if the agent accepts that claim (more precisely; if he accepts that he has more reason to [do x] than to do anything else). This is a basic connection. When the reason is an explanation of his action, then of course it will be, in some form in his [motivational set] S, because certainly- and nobody denies this- what he actually does has to be explained by his S.

Internalist theory explains how it is that the agent’s accepting the truth of ‘There is a reason for you to [do x]’ could lead to his acting, and the reason would thus explain his action. It is obvious on the internalist view how this works. But suppose we take the externalist view, and so accept that it can be true that A has a reason to [do x] without there being any shadow or trace of that presently in his S. What is it that the agent comes to believe when he comes to believe that he has a reason to [do

x]? If he becomes persuaded of this supposedly external truth, so that the reason does then enter his S, what is it that he comes to believe? This question presents a challenge to the externalist theorist?<sup>45</sup>

Most importantly, what's wrong about this argument is that it rests on, what seems to be, a false assumption about normative reasons. Let's take it from the beginning. Williams begins:

The point is the interrelation of explanatory and normative reasons. It must be a mistake to simply separate explanatory and normative reasons.

To this, I reply: 'No'; when we are talking about these different kinds of reasons, we are trying to say completely different things. Both kinds of reasons are, though in distinct ways, explanatory. Normative reasons are explanatory in that they explain why we *should* act in certain ways; they are explanatory in what we can call a justificatory sense. However, they are not, and need not be, explanatory in the sense that explanatory reasons are; they need not explain why we act as we do. The first kind of reasons functions as answers to normative questions, while the second kind answers causal or motivational and/or psychological questions.

Once again I would like to emphasize how Williams seems conflate normative and motivating reasons in a mistaken way. Consider this remark of his:

If something can be a reason for action, then it could be someone's reason for acting on a particular occasion, and it would then figure in an explanation of that action.<sup>46</sup>

This could, depending on how you want to understand it, be accepted by the externalist. When explaining why somebody acted as he did, we could, I believe, say that he acted because he believed that some fact obtained, and this belief could be the reason that made it true that he acted as he should have. This doesn't mean that the reason itself, or the normative truth that something gave a person a reason to act, has any causal power. So, the following need not be considered as being a critique of the externalist position (the question that Williams discusses in the passage I here quote is whether Owen, who doesn't want to join the army, could nevertheless have reason to do so):

Now no external reason statement could *by itself* offer an explanation of anyone's action. Even if it were true (whatever that might turn out to mean) that there was a reason for Owen to join the army, that fact by itself would never explain anything that Owen did, not even his joining the army.<sup>47</sup>

Returning to the passage we are looking at, Williams continues:

If it is true that A has a reason to [do x], then it must be possible that he should [do x] for that reason; and if he does act for that reason, then that reason will be the explanation of his acting.

Here Williams conflates a more plausible claim with the one that I have just rejected. The plausible claim is that our theory must allow for the possibility of us acting for good reasons.<sup>48</sup> But, this is not the same claim as the claim that these normative reasons, in a stronger sense, *explain* why we act. What explains why we act is rather our believing that we have these reasons (or something like that). As we can note, Williams himself have actually said things that go badly together with the claims he here makes. In 'Internal and External Reasons' he says that we can have false beliefs about what reasons we have<sup>49</sup>. More importantly, Williams also says that "[t]he difference between false and true belief on the agent's part cannot alter the *form* of the explanation which will be appropriate to his action"<sup>50</sup> If the agent falsely believes that she has some reason and acts accordingly, then how could this acting be explained by this reason? It couldn't because there is, in this case, no such reason. What explains her acting is rather her being motivated by her belief that she has this reason. In this case we can explain her acting and we don't appeal to any normative reason in doing so. We, instead, appeal to her *belief* about a normative reason. Now, Williams claims that, when explaining why agents act as they do, the form of explanation isn't altered in cases where agents have true or false beliefs. The question, then, is: why should we appeal to normative reasons when explaining agents' actions when these agents act on true normative beliefs? Even if there are no such things as normative reasons, we could, I have claimed, still explain people's actions. And, if people always acted as they have most reason to act, we still wouldn't need to appeal to these agent's normative reasons when explaining their actions, only their beliefs about reasons. So, it doesn't follow, as Williams seems to think, that

the claim that he has a reason to [do x] (...) introduces the possibility of that reason being an explanation; namely if the agent accepts that claim (more precisely; if he accepts that he has more reason to [do x] than to do anything else). This is a basic connection.

Here, again, Williams seems to be confusing the two different claims for being one and the same. On the one hand we have the claim that a normative reason can explain why we act as we do, while, on the other hand, we have the claim that it is the fact that an agent believes that she has reason to act in some way that explains her acting. So, what Williams takes to be a basic connection is *not* a basic connection. We can, I believe, deny the following as well:

When the reason is an explanation of his action, then of course *it* will be, in some form in his S, because certainly- and nobody denies this<sup>51</sup>- what he actually does has to be explained by his S.<sup>52</sup>

Now, it may be true that what an agent does is explained by her S. But that doesn't mean a normative reason has to be in somebody's S in any other 'form' than as the content of a belief. Therefore, the reason itself need not be referred to in explaining the agent's acting. Only the belief (perhaps taken together with some desire) needs to be referred to.

Williams continues:

Internalist theory explains how it is that the agent's accepting the truth of 'There is a reason for you to [do x]' could lead to his acting, and the reason would thus explain his action. It is obvious on the internalist view how this works.

This may not be true either. Although the internalist can explain how an agent coming to believe that she has reason to act in some way may explain her acting accordingly, this again doesn't mean that it is the reason that explains why she acts. To say that somebody has a reason to do something, on the internalist view, is to say that there is a sound deliberative route from this person's S to the conclusion that she should do this thing. What explains this agent doing this thing would then rather be her having deliberated and having decided to act. The deliberation may not have been, in Williams' sense sound, but

this shouldn't alter the form of explanation of the agent's acting. So, if the agent had deliberated soundly, then we would have explained her acting in a similar way. In neither of these explanations do we need to appeal to internal reasons in order to explain the agent's acting.

Williams then turns to considering externalism and writes:

But suppose we take the externalist view, and so accept that it can be true that A has a reason to [do x] without there being any shadow or trace of that presently in his S. What is it that the agent comes to believe when he comes to believe that he has a reason to [do x]? If he becomes persuaded of this supposedly external truth, so that the reason does then enter his S, what is it that he comes to believe? This question presents a challenge to the externalist theorist.

What the agent comes to believe is precisely that he has a good reason to act. It is, I believe, as simple as that. Such a belief could, and/or would, be irreducibly normative. We don't have to believe anything further. This possibility is not one which Williams, at all, seems to have considered. He writes:

*What is it that one comes to believe when one comes to believe that there is reason for him to [do x], if it is not the proposition, or something that entails that proposition, that if he deliberated rationally, he would be motivated to act accordingly?*<sup>53</sup>

Williams here seems to assume that a reductionist answer necessarily is what we are looking for. This should not be presupposed. To answer Williams' question we might say that it's, again, as simple as this: the agent believe that he has a reason to act. That is what it *is* to have a normative belief.

*Acting for a reason*

The argument of Williams just discussed, I have argued, fails. There was, however, one thing that he said that I took to be plausible taken at face value, namely the claim that

If it is true that A has a reason to do x, then it must be possible that he should do x for that reason

We can, following Korsgaard<sup>54</sup>, call this *The internalism requirement*. If we were to accept this claim, would this cause any problems for our externalist view?

Dancy, who defends an externalist view of reasons, thinks so. The problem, he claims, surfaces if we accept a psychologistic<sup>55</sup> theory of motivation/motivating reasons. Dancy's worry is that the combination of externalism and a psychologistic theory "renders us incapable of acting for a good reason"<sup>56</sup>. He writes:

[T]he psychologistic story fails to meet at least two of the main constraints on accounts of motivating reasons. Either it fails to show how it is possible for the reason in light of which we acting to have been among the reasons in favour of acting, or it fails to show how it is possible more generally for a motivating reason to be a good reason for acting.<sup>57</sup>

So, if the fact that jumping out of the window would save my life is a good reason for jumping, this cannot, Dancy claims, on the psychologistic story be a motivating reason since it is a *fact/truth* and not a *belief*. If my motivating reason for jumping was that I believed that it would save my life then this motivating reason and the normative reason are two separate reasons, i.e. the motivating reason is a belief while the normative reason is a truth/fact. A belief about a fact is something else than the fact itself. Therefore it seems that on a psychologistic account of motivating reasons normative reasons cannot be motivating reasons. This would violate the internalism requirement. Thus if we accept the internalism-requirement, externalism and accept psychologism, we here, or so Dancy thinks, have a problem.

What, then, should we make of this? Well, suppose jumping would save my life and that I believe this. Because I believe it I jump. Was the normative reason for me to jump something else than my motivating reason for jumping? That I believe that jumping would save my life is surely something else than the fact that jumping would save me. However, the content of my belief is the same as the fact that jumping would save my life. Let's call this fact 'p'. So when I come to believe that p, then p is the content of my belief and if p is true, then I have a true belief. The normative reason for jumping is the truth of p. Whether I have a good reason to jump does, in other words, depend (in part) on the truth of p. Now,



as Williams noted,<sup>58</sup> it doesn't matter whether our beliefs are true or false when it comes to their ability to motivate us. False beliefs motivates just as true beliefs do. It is also the case that our *rational* ability to reason sometimes makes us *irrationally* jump to conclusions, which might make us act on false assumptions. Motivating and normative reasons differ, I have said, in the sense that when I act intentionally we can always say that there is a motivating reason explaining my acting, whereas there might not be any *good* reason to act in such a way. This suggests that whether I have a normative reason to act in a certain way depends on something else than whether I have a motivating reason for acting when acting. The condition for there to be a motivating reason, I have said, is that I act intentionally. Differently put, we can say that whenever I act intentionally there is some motivating reason explaining my act. Whether I have a good reason to act depends, roughly, on facts about the given situation. If, for example, Paul is being attacked by a tiger he has a good reason to run.

What I think that we need to do is to question the authority of the internalism requirement. Or, more precisely, we should, if we wish to keep it, see if there is a better way of formulating it. My suggestion is that it should read:

*The improved internalism requirement:* Something cannot be a normative reason unless it can be the content of a belief that is capable of being, or being a part of, an explanatory reason for someone to act.

If we allow this change then the problem appears to be solved. Remember again that consideration of normative reasons and of motivating reasons are two different things in the sense that when we are speaking of good reasons to act we are not trying to explain why someone *would or might* act in a certain way; we are trying to explain why this person *should* act in the given way. It is only, I would say, if normative and motivating reason are incapable of being related in *any* plausible way that we seriously need to doubt whether it makes sense to speak of normative reasons as something distinct from motivating reasons.

Remember what I, at the beginning, said about the difference between explanatory and motivating reasons. When we speak of explanatory reasons we are trying to explain why someone acted as he/she did. The most common way to do this is to appeal to the agent's motivating reasons. When we ask why Paul is running we will most likely expect to hear, as an answer, Paul's motivating reason, namely that a tiger is chasing him. So from Paul's perspective the reason that he is running is that a tiger is chasing him; because he believes

that this is true he runs. Now, if a tiger is chasing you, you do seem to have a *good* reason to run, so Paul is acting for what seems to be a good reason. It is, then, the content of his belief that he is being chased by a tiger that motivates him. What's important here is the *content* of his belief rather than the fact that he believes something, even if there here is a dependency-relation. Had he not believed that he is chased by a tiger, he would not have been moved to act<sup>59</sup>.

You can act for a good reason since the fact that you believe that you have this reason can motivate you to act. What explains why you acted in certain way might be that you had a certain belief. What motivated you was the content of this belief and if you didn't falsely believe that you had a good reason to act, then we could say that you acted for a good reason.

### *Why reasons for having desires might be more important than reasons for acting*

Reasons for having certain desires are, I believe, of greater importance than reasons for acting. The same might be true for reasons for believing as well. Why is this so? Suppose that we think that the internalist theory of reasons for actions is true. Reasons for actions are all, on this view, provided by desires. If we assume that this is a correct description of having reasons for acting, it is still perfectly reasonable to ask whether we have any reasons to have these desires. We could, in other words, believe that the internalist-model gives a correct analysis of reasons for *acting*, and still think that this doesn't matter much. What is most important, we could say, is the question of whether we have any reason to have the desires that would be fulfilled if we were to act in certain ways. Those who accept the belief/desire-theory of intentional action and therefore believe that desire is a necessary condition for action should certainly accept this.

Now, remember that internalism is a development of Hume's theory of practical reason. The Humean believes that desires, or 'passions' in Hume's terminology, cannot be contrary to reason. The internalist doesn't have to say this. There can be desire-based reasons for having certain desires. Desires that we have because we have another, more fundamental, desire can be called 'instrumental' desires. There can, as we can note, be long chains of instrumental desires. We should also, and more importantly, note that these chains must end somewhere. At the end of the chain there has to be a desire for something for its own sake. Such a desire would be an 'intrinsic' desire, or a 'final' desire, and here

the internalist has to stay true to his Humean roots. He must, on internalism, claim that these desires are desires that we have no reason to have or not to have.

The internalist may object to what I say here. There can be desire-based reasons for having intrinsic desires, he might say, and point our attention to, for instance, the, in our discussion already mentioned, paradox of hedonism. This, we said, was the claim that, though pleasure is the only thing that actually is of value, we ought to desire other things than pleasure for their own sake, for if we don't, we will be less successful when trying to achieve our hedonistic aim. Since we have an intrinsic desire for pleasure, we have reason to do what is most likely to bring us pleasure and that is to intrinsically desire *other* things than pleasure. Hence, the internalist may conclude, we can, on internalism, have reasons to have intrinsic desires. But these intrinsic desires are not intrinsic desires in the most important sense. I take the most relevant intrinsic desires to be those desires that are ultimately grounded in the desirable features in the object(s) of desire. These are the desires that the internalist cannot claim that we have any reason to have. The reason for us to desire pleasure is, we can note, exactly such an example.

Some would claim that the only thing we have intrinsic desires for is pleasure and/or happiness. I believe that we should reject this claim. It is perhaps true of some people, but many of us desire other things than happiness for their own sake. Examples of things desired for their own sake could include knowledge, a morally commendable lifestyle, friendship, excellence in arts or sports, posthumous fame, beauty, and so on. Whether these particular things actually are desirable or valuable for their own sake is something that I will not discuss here, even though I believe that some or all of these things are. Neither will I argue here for the claim that we *desire* other things than happiness for their own sake<sup>60</sup>.

Now, as we can note, it appears counter-intuitive to think that at the end of our chains of desires there are always desires that we have no reasons to have. As Parfit points out, even Hume seems to grant this in a passage of the *Treatise* where he seems to forget his own theory. In this passage, Hume writes:

Ask a man *why he uses exercises*; he will answer *because he desires to keep his health*. If you then enquire *why he desires health*, he will readily reply *because sickness is painful*. If you push your enquires further and desire a reason *why he hates pain*, it is impossible that he can ever give any. *This is an ultimate end*, and is never referred to any other object (...) beyond this it is an

absurdity to ask for a reason. It is impossible there can be a progress *in infinitum*; and that one thing can always be a reason why another is desired. Something must be desirable on its own account.<sup>61</sup>

Parfit responds to this in the following way:

For 'desirable' Hume should have written 'desired'. Something is desirable if it has features that give us reasons to want this thing. Hume denied that there can be such reasons.<sup>62</sup>

What I have claimed is this: even if it would be true that a reason to *act* in a certain way is provided by the fact that this act would help fulfil a desire one has, it would still be very much possible, and important, to ask why one should have this desire. This question will be especially important if we accept the belief/desire-theory of intentional action. To answer by citing a further desire will be unsatisfactory because at every chain of instrumental desires there will be an intrinsic desire. Having a certain desire cannot, I believe, give you a reason to have any other desires in any direct way.

If you desire a certain end, you might have reason to desire the means, but this reason would not be provided by the first desire. This reason would rather be provided by the fact that the end is worth achieving, or that desiring the means would make you more successful in pursuing the end. Consider the following example. You desire to achieve excellence in some sport. The proper means to this could be working out and going to practice several days a week. Feeling no desire to work out and practice will certainly cause nuisance. It can, of course, be questioned whether you have any reason to desire the end in question. If there is such a reason, it is most certainly not provided by a further desire.

If there is no reason to pursue a certain end, then how could my desiring this end provide me with a reason to desire the means to the end? If it would be better if I didn't desire the end, then it seems that it also would be better if I don't desire the means. Imagine that you know that there are reasons against achieving a certain end that you desire. You strongly desire to experience the effects of a drug you are addicted to. The voice of reason tells you that achieving this would be bad for you, as this drug messes up your sense of reality. But, since you are addicted, you strongly desire to take the drug, which is the means to the end, i.e. to get into this state of euphoric delusion. Do you have

reason to the desire the means to your desired end? The answer seems to be ‘No’. Contrast this with an example in which you’re ill and you are thinking about taking a certain drug because you believe that it will help you achieve the end of getting you well again. You here desire the end and you seem to have reason to desire the means. Doesn’t it appear to be the fact that you will get well, rather than the fact that you desire this end, that gives you reason to desire the means?

Some would say that we should, or perhaps always do, necessarily desire the means to the ends we desire. This is a little too strong. Achieving a certain end might be desirable, but since the sacrifice involved in pursuing the end would be too great, you might not desire to take the proper means to your end, and, besides, you might actually have reasons to abstain from pursuing the end.

### *Rational aims and intentions*

Our aims are rational insofar as they are supported by reasons, and if our aims are rational we have reason to do what will help achieve these aims. But, we have these reasons because these aims are rational, not simply in virtue of our having the aims in question. We cannot, in that way, create reasons; we cannot merely decide what reasons we have. But, our decisions and/or intentions may have consequences that ‘create’ reasons. Because I intend to go on a trip, I buy tickets. Now that I have bought the tickets I have reason to use them. If I don’t I shall have wasted my money. But here it wasn’t the intention that was the reason. One cannot, in Michael Bratman’s terms, ‘bootstrap’<sup>63</sup> reasons into existence. If we could, then I could, just by intending it, suddenly have a reason to steal your wallet. If you’d complain, I’d just say that I had a reason to steal it—I intended to. It’s obvious that you need not accept this.

### *Reasons, beliefs and the appeal to consistency*

Suppose now that you thought you had a good reason to act and therefore acted, when it really wasn’t the case that you had this reason. If you have false beliefs it may be rational to do what you have no reason to do. This does, however, depend on which kind of belief of yours that is false. If you have false your *non*-normative beliefs, this may not render your acting irrational. Thus if you falsely believe that the room is on fire, you are acting rationally if you leave it. You had an apparent reason and you responded to it. But your

belief that the room is on fire may be irrational though. If you feel no heat, see no flames or smell no smoke, you have no reason to believe that the room is on fire and your belief is irrational. But as the building you are in is on fire, it may be good that you believe that your room, which is not yet on fire, is in flames. Believing this has good consequences in that it makes you jump, which is something you have reason to do. Does it mean that you *did* have reason to believe what there was no apparent reason to believe? If it does, it cannot, though, make it true that your belief was rational. I shall return to this issue later in this section. First I will say a little about beliefs as such.

Beliefs can be either non-normative or normative<sup>64</sup>. Non-normative beliefs are beliefs about what is the case. Normative beliefs are beliefs about what ought to be the case, i.e. beliefs about what we ought to do, or want. What about beliefs about values? One way of understanding such beliefs is this: If I believe that something is of value, then I believe that there are reasons to care<sup>65</sup> about this thing that are provided by features, or properties, that the thing in question has. If, for instance, I believe that living is good, then I believe that there are reasons for me to want to live. These, I am inclined to think, are basically two different ways of reporting the same belief. Compare:

(V) Suffering is bad

with

(N) We have reason to avoid suffering

Since these two claims do not appear to have different truth-makers, and since no new information appears to be given in (N) that is not given in (V), we need here not have two different discussions about these kinds of beliefs. I will therefore here only talk about normative beliefs, as I think that evaluative beliefs can be understood in terms of normative beliefs.<sup>66</sup>

We can note that particular normative beliefs always depend on non-normative beliefs in the end. Believing, for example, that I ought to jump out of the window in the example I called *The fire* ‘depends’ on the belief that I would die if I were to stay in my hotel-room. That you most likely will be killed if you’re caught by the flames is not a normative belief. It might be objected that some of our beliefs about what we ought to do depend instead on other normative beliefs. If I believe that I ought not to kill people, I might have this normative belief because I have the normative belief that killing is wrong. But believing that I ought not to kill anyone and believing that killing is wrong might be considered as

being the same thing. Better put we might say that that believing that murder is wrong implies that you believe that you ought not to kill anyone (at least under normal circumstances). Anyway, the belief that killing is wrong still has to depend on non-normative beliefs. Something isn't just wrong; there is a dependency-relation between normative properties of things or acts and their natural, non-normative properties.

This is not to say that normative truths are reducible to non-normative truths though. The possible truth of 'x causes pain' is another kind of truth than the possible truth of 'since x causes pain, we have a reason to avoid x'. Believing that x causes pain seems to give us a reason to believe that we ought to avoid x, but it doesn't necessarily imply that anyone who believes that x will cause pain will believe that she ought to avoid x. The normative belief is separate from the non-normative. It does, however, depend on the non-normative belief. We take, I have said, some facts to have normative significance. If we are taking them to have such significance, these facts matter and are, thus, reasons.

What, then, makes beliefs rational or irrational? The most obvious answer is the one already given: beliefs are rational when they are formed as a 'response' to apparent reasons to believe—there is evidence counting in favour of believing something. This answer appeals to the origin of the beliefs. Thus if Jim's fingerprints are on the murder-weapon, we seem to have reason to believe that he is the murderer. Suppose though that he actually is innocent. Are his fingerprints on the weapon still a reason for us to believe that he is guilty of this crime? It certainly seems so. It is an apparent reason to believe that he is the murderer and it makes believing that he is rational. So it seems that it can be rational to believe what is false if there are apparent reasons to believe. The truth in this case, we have said, is that Jim is innocent, but, given that the evidence seems to point in the opposite direction, it appears that believing the truth, in this case, would be irrational. Thus it can be rational to believe what is false. These kinds of evidential reasons for believing do not make the rationality of our beliefs truth-relative. After having looked at another way in which the origin of our beliefs can make our beliefs rational or irrational, we will consider intrinsically rational beliefs and I will then say more about the relation between what is true and our beliefs.

Now, another possible origin of beliefs is processes of theoretical reasoning, from which we can draw conclusions about what we take it to be rational to believe. We can call this 'The appeal to consistency' as this has to do with how our beliefs go together with each other. If someone believes that the world was made in six days, it follows that this person, for the sake of consistency, ought rationally to believe that the world was made in

less than a week. Denying one of these claims while embracing the other would clearly be a rational failing. Does this mean that the person who believes the first claim, from which the other follows, has *reason* to believe that the world was made in less than a week? John Broome, from whom I have borrowed this example, certainly doesn't think so. He writes:

No one ought to believe that the world was made in less than a week; the evidence is strongly against it. Even if you believe that the world was made in six days, still it is not the case that you ought to believe it was made in less than a week. Nevertheless, you ought (to believe that the world was made in less than a week, if you believe in six days). You can satisfy this requirement either by not believing that the world was made in six days or by believing that it was made in less than a week. As it happens, you ought to satisfy it the first way. You ought not to believe that the world was made in six days, even if you do.<sup>67</sup>

Broome notes that there is no apparent reason, i.e. no evidence that counts in favour of believing that the world was made in six days, or less than a week. So you ought not to believe it; believing it would be irrational. But still, he claims, it may, in another sense, be the case that you ought to believe that the world was made in less than a week. What does Broome mean by this?

His point is this: normativity is not all about reasons. If we think so, we will overlook other important features being rational involves and falsely take things that aren't reasons to be reasons. If you ought to do, desire or believe something, it is often the case that there is reason to do so, but that is not always true. There is, Broome claims, another sense of 'ought' that shouldn't be confused with the sense just mentioned. This later sense in which there may be 'oughts' has to do with the requirements of correct reasoning, theoretical or practical. You are, in Broome's terms, *normatively required* to believe what follows from your other beliefs. What Broome is talking about here is basically what I have called the appeal to consistency.

There may arise some confusion here. Though Broome's point is one I believe that we should accept, we can ask if it isn't true that we have reason to accept that there are such normative requirements that Broome talks about. If we have such a reason, i.e. we have reason to believe that we ought to believe what follows from our other beliefs, then it may seem that we really do have reason, in the case introduced above, to believe that the world



was made in less than a week. For if we have reason to accept the idea of normative requirements, then wouldn't that mean that we have reason to believe the consequences of constraining our beliefs in accordance with this idea of consistency? If we, in other words, have reason to have consistent beliefs, doesn't that mean that the person in the example has *reason* to believe that the world was made in less than a week?

Things, I believe, do, here, get a bit tricky. We can, in order to sort this out, distinguish reasons for accepting belief policies from reasons to believe particular propositions. If our beliefs aren't consistent, this creates problems for us. That having consistent beliefs could have good consequences may, in other words, be a reason to accept the idea of normative requirements on our reasoning and beliefs. So we have, in this sense, reason to believe what follows from our other beliefs. But, as these other beliefs may be irrational, it is hard to believe that we have reason to believe some of the things that may follow from them. This may lead us to think that we only have reason to believe what follows from rational beliefs, but to think this would be to confuse reasons to accept belief policies with reasons to believe particular propositions. Compare this with how it can be rational to do what one has no reason to do if our belief that we have this reason depends on false non-normative beliefs. That we are being rational in such cases depends on our responding to apparent reasons. This, however, does not seem to be the case when it comes to beliefs that follow from other beliefs. That you believe that the world was made in less the six days is not an apparent reason to believe that it was made in less than a week. So the analogy breaks down here, but the point remains that it may be rational both to do and believe what one has no reason to do or believe. To think that beliefs could be reasons for having other beliefs would be to think that we, in a sense similar to the one rejected in an earlier section, could 'bootstrap' reasons into existence; I start believing p and suddenly I have a reason that I didn't have before to believe q.

### *Intrinsically rational beliefs and instrumentally rational beliefs*

Though the rationality of beliefs does to a large extent depend on the origins of the beliefs, this, I'm inclined to think, isn't the only thing that could make beliefs rational, or irrational. When I talked about desires I said that it may be that desires can be intrinsically rational or irrational. What I then meant was that these desires were rational in virtue of the desirability of the objects of desire. If, to put it differently, something has features that give us reasons to want this thing for its own sake, then our desiring it would be intrinsically

rational. As I believe, beliefs, too, can be intrinsically rational or irrational. While what made desires intrinsically rational is facts about the object of desire, it is the content of a belief that can make that belief intrinsically rational, or perhaps sometimes more obviously, intrinsically irrational. On an internalist view this can, obviously, not be true when it comes to normative beliefs, but on the realist view that I am trying to defend it can. This, I think, is a strength that the realist position has. Suppose that somebody has no false beliefs about the nature of pain. Though this person knows how pain feels and what it is to suffer, this person believes that it is rational to want pain and suffering for its own sake. This, it may be objected, is an incredible example. Who, we might ask, would ever have such a belief? But this is no objection as it only proves that most of us are inclined to believe that such a belief is so irrational that no rational person could have it. This belief would be intrinsically irrational. If a belief is intrinsically rational, then it is the content of that belief that makes it so.

Suppose now that you believe something to be true. Does this make it intrinsically rational to believe that which you take to be true? If I, for instance, believe it to be true that the earth is flat, would failing to believe this be intrinsically irrational? I am now assuming that it is possible to believe something to be true without believing it. This may not be possible, but the point here is not to discuss that matter but to ask whether beliefs about the truth of beliefs one has have any relevance to the rationality of these beliefs. Some of our beliefs are such that we don't have any strong second-order beliefs about whether these first-order beliefs are true or false. If I believe that the weather is going to be nice next week, I might not feel that I have any good reason to believe that this belief of mine is true. Anyway, if such a thing is possible, it appears irrational to believe something to be true and still fail to believe it. This is not to say that a belief to the effect that some other belief  $p$  of ours is true always gives us reason to believe  $p$ . Rather, it is, at least in the case of beliefs about empirical facts, that which makes us believe that what we believe is true that gives us reason to believe that this belief of ours, whatever it may be, is true. This may appear to be a trivial point, but what I am trying to say is, basically, the following: Second-order beliefs about the truth or falsity of our beliefs have impact on whether these first-order beliefs are rational or not, but they are not, in the case of non-normative beliefs, reasons for believing or not believing the first-order beliefs in question.

Moreover, what is more important is this: whether a second-order belief about the truth of a first-order belief makes the first-order belief intrinsically rational depends on what this first-order belief is about and whether there can be better ways of finding out whether your

belief is rational. If our first-order belief is a belief about some empirical fact, then we can, instead of asking whether we believe that our belief is true, ask if there are any apparent reasons for believing that this fact obtains. But, when it comes to other kinds of beliefs, such as beliefs about values, normative truths, logical truths and so on, it may not be the case that we can test our beliefs in any other way than by asking ourselves whether these beliefs appear to be true or not. We can, knowing all there is to know about pleasure, ask, without making any linguistic or logical mistake, whether it is something that it is rational to want for its own sake. Does the fact that the answer clearly seems to be ‘Yes’ give us reason to take it to be rational to believe this? My answer is ‘Yes’; that this obviously is the right answer makes it intrinsically rational to believe this. That something seems *obviously true* can also, I believe, in these cases be said to be at least an apparent *reason* to believe it.

There can also be reasons for believing some things that are provided by the consequences of believing. These reasons are reasons that we may be never aware of. But that, I believe, is not a problem. If we knew that it would be good for us to believe something that we have no apparent reason to believe we might, at best, start hoping that this, whatever it may be, was true. It is hard to believe that rational individuals could cause themselves to believe what there seems to be no apparent *evidential* reason to believe. Causing, if possible, oneself to believe such a thing would be, what we might call, *rational self-deception*. If I, for instance, enter a cave and the only exit that I believe the cave to have suddenly becomes blocked, I may not take myself to have any evidential reasons to believe that there is any other exit; it’s dark and, to the best of knowledge, the now blocked exist was the only way out. This may cause despair on my part. If I thought that there was some other way out, I would probably be more motivated to look for it. Thus I would, most likely, also be more successful when trying to do so. So, it seems that it would have good consequences if I believed that there was another way out. Trying to cause myself to believe this, therefore, appears rational. If having some belief has good consequences, then one, I believe, has reason for having this belief. We can call such a reason an *instrumental* reason for believing, and we can say that such beliefs can be *instrumentally* rational.

We have, then, four ways in which beliefs may be rational. Rational beliefs can be:

Reason-supported beliefs:	The evidence points in favour of believing
Consistency-required beliefs:	Beliefs that follow from other beliefs we have
Intrinsically rational beliefs:	The content of the belief appears to give us reason to have this belief, or the belief appears obviously true
Instrumentally rational beliefs:	The effects of having the belief give us reasons to have it

With this in mind we can turn our attention back to desires.

### *Rational irrationality*

As I have already said, we desire some things for the sake of their effects, i.e. we desire some things because they would help us achieve something else. While these desires are instrumental, there are also intrinsic desires. We want certain things for their own sake. Such desires are intrinsically rational if there are reasons to have them. There might, however, be desires that, although they are intrinsically irrational, we still have reason to have. When discussing this, Parfit writes:

If we believe that having some desire would have good effects, what that belief makes rational is not that desire itself, but our wanting and trying to have it. Irrational desires may have good effects. Thus, if I knew that I shall be tortured tomorrow, it might be better for me if I wanted to be tortured, since I would then happily look forward to what lies ahead. But this would not make my

desire rational. It is irrational to want, for its own sake, to be tortured. The good effects of such a desire might make it rational for me, if I could, to cause myself to have it. But that would be a case of rational irrationality.<sup>68</sup>

It is, as Parfit writes, clearly intrinsically irrational to want to be tortured. But, since having such a desire would have good effects, it seems rational to, if possible, cause myself to have it. Thus, according to Parfit, there can be desires that it is rational to want and try to have, but irrational to have. Wlodek Rabinowicz and Toni Rønnow-Rasmussen disagree and write:

While attractive, this line of thought is unconvincing. To be sure, if a pro-attitude is either instrumentally or finally valuable, we do have reasons to want it or to try to have it. (...) But why should we deny, as Gibbard<sup>69</sup> and Parfit want us to, that we also have reasons to *have* these valuable attitudes? (...) As far as we can see, (...) Parfit [has] no argument for [his] claim, apart from appeal to an intuition that we don't share.<sup>70</sup>

The views of both parties, I believe, make sense. In one sense it seems obvious that I, in Parfit's example, have a reason to have a desire to be tortured. Does this mean that it can be rational to want to be tortured for its own sake? Here the answer seems to be 'No'. What should we think of this? Compare this with the case we discussed earlier where it was good for me to believe what it would be irrational to believe. It would have been good if I believed what there was no evidential reason for me to believe. But since I, as a rational individual, find it hard to believe what there seems to be no reason to believe, the thing to do seemed to be to try to, in some way, deceive myself. Such self-deception would, in this case, be rational, since it would bring about good effects. The same seems to be true in the case where wanting to be tortured would have good effects. If I am (substantively) rational it seems hard to imagine that I ever would desire to be tortured, but it is imaginable that I could realise that it would be better if I did have this desire. Thus, it would be better in this case if I was irrational, and therefore the rational thing to do seems to be to try to become irrational.

While these considerations seem to favour Parfit's view, they also seem to favour the view that we can have reasons to have irrational desires and beliefs. We should, I think,

agree with Parfit that these cases are cases of rational irrationality. Believing that it would be rational to want to be tortured cannot make it rational to want to be tortured. But since being irrational could have good effects, it can be rational to cause oneself to become irrational. If I already have a desire to be tortured, I would be deeply irrational. But, if I don't have this irrational desire, I could, rational as I am, realise that it would be better if I was irrational during a certain period of time, and, therefore cause myself to become irrational if that is possible. This would be rational self-deception or, as Parfit calls it, rational irrationality. So although there *can* be reasons for having irrational beliefs or desires, this doesn't mean that such beliefs or desires would be intrinsically rational. This, to me, does not appear strange.

We can, in order to illustrate this further, consider these four ways<sup>71</sup> in which desires can be supported by reasons and thereby be rational:

Intrinsically rational desires:	Facts <sup>72</sup> about the object of desire give us reasons to have this desire
Instrumentally rational desires:	The effects of having a desire give us reasons to have this desire
Intrinsically rational but instrumentally irrational desires:	Though facts about the object of desire give us reasons to have this desire, the effects of having the desire give us reasons not to have it
Intrinsically irrational but instrumentally rational desires:	Though facts about the object of desire give us reasons not to have this desire, the effects of having this desire give us reasons to have it

A desire like the desire in Parfit's example to be tortured is intrinsically irrational but instrumentally rational, and therefore we can, I believe, say that there can be reason for, not only wanting and trying to have it, but also for actually *having* it. And as this desire is, in one sense, irrational, while it, in another sense, is rational, the label 'rational irrationality' seems fitting.

We can also consider an example in which a desire is intrinsically rational but instrumentally irrational. Suppose I live in a dictatorship and that, as I'm not too happy about it, I desire to live in a democracy without an evil dictator. Now if this desire becomes known, I may be in big trouble. So here we have a case in which my desire, although it is intrinsically rational, is instrumentally irrational. This, then, becomes a case of what we can call 'irrational rationality'.

Remember now that I said that claims like 'x is good' imply that there are reasons to care about, or want, x. This should *not* be taken to imply that being tortured is good. That would only have been true if wanting to be tortured was intrinsically rational. It is *wanting to be tortured* that, in this example, is instrumentally good. So whether it is the object of the desire or the desire itself that is good depends on whether the desire is intrinsically or instrumentally rational. If the desire is both intrinsically and instrumentally rational, then both the object of desire and the desire itself are good. If we accept this we will, I believe, dodge objections that, by means of counter-examples, try to show that it cannot be true that to say that something is good is basically the same as saying that we have reason to care about this thing.<sup>73</sup>

*Does 'ought' imply 'can' when it comes to reasons?*

It may now be asked how it could be that I, in the example where I am waiting to be tortured, can have a reason to *have* a desire to want to be tortured if I find it extremely hard to start wanting this. If we do assume that it, by means of sheer willpower, is virtually impossible to go from not wanting to be tortured to wanting to be tortured, then we might, from this, be tempted to draw the conclusion that one cannot have a reason to have this desire. One, we might say, cannot have reason to want something if it is impossible—or, at least, extremely hard—to want this thing. I do, however, believe that this would be a mistake. When it comes to reasons the principle that 'ought' implies 'can' does not apply.

This is part of the point of the externalist view; thinking otherwise would, I think, be to take the internalist view. On the other hand, the principle in question is true when it comes to rationality judgments. Think of the example where my hotel is on fire. Suppose that my dog is with me in my room. Like me, it will, if it stays in the room, be badly burned or die. This makes it true that my dog, just I do, has a reason to jump out of the window along with me. My dog is not generally rational like me though. It could not, like me, understand that it has a reason to jump in the sense that I can. You cannot tell my dog that, since the room will soon be on fire, it has reason to jump out of the window. Nevertheless, it could still act like it has reason to act; it will most likely flee fire. However, it does not appear meaningful to call my dog's behaviour rational or irrational. It lacks the rational ability that persons have to think in terms of reasons. Beings that aren't generally rational cannot act rationally or irrationally. That is something that only we, as persons, can do.

The externalist, as I have said, doesn't believe that it is our actual or counter-factual desires that provide us with reasons. Therefore, we don't have to claim that 'ought' implies 'can' when it comes to reasons. There is, thus, an asymmetry-relation between reasons and rationality in the sense that, while what one has reason to do isn't relative to whether one could, or would, be motivated by the belief that one has such a reason, it is only those who are generally rational that can rightfully be called irrational at times. Rationality is here, then, partly defined in terms of ability to respond to reasons, but reasons are not defined in terms of rationality. Taking such a non-reductionist view on reasons lets us avoid charges of circularity. Rational beings, we can say, are beings that believe that some things matter and that they should care about these things and live in certain ways. But, we need not, for this reason, disallow the possibility of these rational beings being mistaken at times. As rational beings, we know, as I said earlier, this to be true; it is part of our rational ability to be able to realise that some of our choices are bad choices and that some of our desires are ones that we shouldn't act on.

*What do reasons do and why are there reasons?*

What, then, is it for something to matter? I shall here not go thoroughly enough into this issue, but only make some remarks about what I take the realist view to imply. We can start with a somewhat controversial metaphysical claim, namely the claim that if something matters, then it has the normative *property* of mattering. What kind of properties, it may now be asked, are these supposed normative properties? Are they natural



properties? If they had been, then, it seems, we could prove that some things do really matter, or that some things really do have these properties. Why? Because natural properties usually (or always) have causal powers, which make it possible for us to, by means of experiments and observations, prove claims we make about them. Now, normative properties, as I've argued, aren't natural properties. They are non-natural properties. I recognise that claims like this usually give rise to objections about metaphysical queerness as well corresponding epistemic objections; if there are irreducibly normative truths, then how do we verify that our beliefs about them are true? Such objections, I believe, can be answered, but I shall not attempt to do so here.

Now, normative properties have no causal powers. This, I believe, is because normative properties are what might be called 'true-about' properties<sup>74</sup>. Suffering has the normative property of it being true about it that it is worth avoiding. Truths have no causal powers. How, it may now here be objected, can it be that these truths matter if they lack causal powers? What do reasons do? What is the point of their existence? Well, reasons do not 'do' anything by themselves *to* us. Reasons do not push us to act in certain ways.<sup>75</sup> What reasons do is rather to make it *true* that we *should* act in certain ways. This is the normative relation between reasons and our motivations that I've been trying to argue in favour of.

Williams believes, as we have seen, that normative reasons can be explanatory reasons for people's acting. I have denied this and said that what reasons explain is, rather, why we *should* live in certain ways. That, if anything, is what reasons 'do'. If we ask how reasons can matter, we would, I believe, be asking whether it matters that some things matter. If that is a real question, its answer would, either way you go, be paradoxical, it seems, especially if the answer is 'No'.

I don't think that we should look for a causal explanation of why there are reasons or normative truths. Compare this with how the existence of the universe is to be explained; is there a causal explanation of why the universe exists? The answer here, I believe, is 'No'; the universe just happens to exist. It could have been the case, though it isn't, that it didn't exist. Though we may find it hard to imagine what that would have been like, we can understand the possibility of it being true. I have taken seriously the possibility of there being no normative truths and that nothing, thus, matters. But, I find it hard to see how it could be true that something like suffering isn't worth avoiding or that the well-being of others is not something that we have reason to care about. And, these truths, I'm inclined to believe, aren't made true by anything external to them; there is no causal explanation to

be given here. So, though one could say that it could have been true that nothing is actually worth caring about, it is hard to imagine what that would be like, in the same way as it is hard to imagine what it would have been like if the universe did not exist. So at the same time as I think that there is an epistemic possibility-- it is possible for all we know--that there are no normative truths, I think that if there are normative truths, then these truths are, or must be, metaphysically necessary truths. By this I, for instance, mean that suffering couldn't have been anything but bad; an opposite claim appears unintelligible. This means that, in all possible worlds, it is true that suffering is bad, or worth avoiding.

It may now be asked if these normative truths are mind-independent. To answer this, we should reflect on what it is that makes things worth wanting, or worth avoiding. On realism the answer, as I've said, is that it is features of the objects of desire, and not our desires that give us reasons to act, or want to act. If there were no beings that could want things or that could reason, then there wouldn't be any beings of whom it would be true that they have reason to want certain things or who could be rational in wanting certain things. Does this mean that normative truths are mind-dependent? The answer, I believe, is 'No'; independently of the existence of rational beings, there could still be normative truths, just like there would be truths about other things. It was true that the universe existed long before there were any rational beings that could grasp this truth. So, if all humans vanished from the face of the earth, it could still be true that if there were any human beings, they would have had reasons to want certain things and to act in certain ways. If there came along rational aliens, or if some other earthly species evolved in a manner that gave them a rational ability similar to, or superior to, ours, then these beings could, I believe, understand these normative truths just as we do. Perhaps they would have a better ability than ours to do so. Some normative question and some questions about values are really hard for us to answer. These questions might be better understood by others.

### *Summary*

I have distinguished between three kinds of reasons; normative, motivating and explanatory reasons. When we talk about these different kinds of reasons, we need to keep in mind that we are talking about different things. There is a deep difference between claims about normative reasons and claims about motivating or explanatory reasons. The difference is, simply put, that normative claims are normative and that they cannot,

therefore, be reduced to claims about motivation or to other causal explanations. Williams' internalism, I have argued, fails to recognise this difference. Internalists claim that there is a necessary connection between our reasons and our motivations. We can, I have argued, deny that there is such a necessary connection, because if we allow it, we have to draw absurd conclusions about what reasons certain people have and, more importantly, don't have.

The view I have been trying to defend, I have called Normative Realism. On this view it is irreducibly true that some things are worth caring about. Though I haven't intended to defend any substantive claims about exactly what we should care about, I have used the example of suffering as something that there is reason to want to avoid. That this seems obviously true seems to be enough for there to be a reason to believe this. Beliefs of this kind are, or so I've argued, intrinsically rational. And this is one out of four ways in which beliefs can be rational. Beliefs can also be rational because they are supported by evidential reasons, because they are consistent with other beliefs that we have, or because having them would have good consequences. I have also distinguished between four kinds of rational desires. Desires can be intrinsically rational, instrumentally rational, intrinsically rational but instrumentally irrational or intrinsically irrational but instrumentally rational. Something that speaks in favour of the realist view is that it allows for the possibility of desires being intrinsically rational.

My main aim in this paper has been to try to explain why I believe that we ought to reject Williams' internalist view on reasons. For this reason I haven't tried to defend normative realism against some of the objections that can be raised against it. As I've said, I do, however, believe that it is possible to do so.<sup>76</sup>

#### *References:*

Bratman, Michael: *Intentions, Plans and Practical Reason*, Cambridge: Harvard University Press, 1987

Broome, John: *Reasons in Reason and Value: Essays on the Moral Philosophy of Joseph Raz*, eds. R. Jay Wallace, Michael Smith, Samuel Scheffler, and Philip Pettit, Oxford University Press

<http://users.ox.ac.uk/~sfop0060/pdf/reasons.pdf>

- Dancy, Jonathan: *Practical Reality*, Oxford: Oxford University Press, 2000
- Gibbard, Allan: *Wise choices, Apt feelings* Oxford: Clarendon Press, 1990
- Hare, R. M. : *Moral thinking* Oxford: Clarendon Press, 1981
- Kagan, Shelly: *The limits of morality* Oxford: Oxford University Press, 1989
- Korsgaard, Christine: *Skepticism about Practical Reason* p. 317 in *Creating the Kingdom of Ends* New York: Cambridge University Press, 1995
- Korsgaard, Christine: *The Meta-Physical Foundation of Normativity*,  
<http://www.people.fas.harvard.edu/~korsgaard/Korsgaard.LL1.pdf>
- Moore, G. E. : *Principia Ethica*, Cambridge: Cambridge University Press, 1993
- Nagel, Thomas: *The possibility of altruism* Oxford: Clarendon Press, 1970
- Parfit, Derek: *Rediscovering Reasons*, forthcoming
- Parfit, Derek: *Rationality and Reasons* in *Exploring Practical Philosophy*, eds. Egonsson, Dan et al Ashgate, 2001 p. 25
- Rabinowicz, Wlodek and Rønnow-Rasmussen, Toni: *The Strike of the Demon: On Fitting Pro-attitudes and Value*, forthcoming
- Rawls, John: *A Theory of Justice*, Cambridge: Harvard University Press, 1971
- Scanlon, T. M. : *What we owe to each other* Cambridge: Harvard University Press, 1998
- Sidgwick, Henry: *The Methods of Ethics*. Cambridge: Hacket Publishing, 1981
- Williams, Bernard: *Internal and External Reasons* in *Moral Luck* Cambridge: Cambridge University Press, 1981

Williams, Bernard: *Internalism and the obscurity of blame in Making Sense of Humanity*  
Cambridge: Cambridge University Press, 1995

Williams, Bernard: *Postscript in Varieties of practical reason* ed. Millgram, Elijah  
Cambridge: The MIT Press, 2001

Williams, Bernard: *Values, Reasons and the Theory of Persuasion in Ethics, Rationality  
and Economic Behaviour* eds. Farina, Francesco, Hahn, Franic and Vannucci, Stefano,  
Oxford: Clarendon Press, 1996

---

<sup>1</sup> Most importantly in *Internal and External Reasons* (henceforth *IER*) in *Moral Luck* Cambridge: Cambridge University Press, 1981, but also in *Internalism and the obscurity of blame (IROB)* in *Making Sense of Humanity* Cambridge: Cambridge University Press, 1995 and other articles.

<sup>2</sup> Constructivism is, very roughly, the view that we have reasons to want or do certain things because there are correct ways of arriving at the conclusion that we have these reasons. Rawls theory with his idea of a reflective equilibrium is an example of this. See his *A Theory of Justice* Cambridge: Harvard University Press, 1971.

<sup>3</sup> I hope to do both of these things elsewhere. Why I think that we should prefer a realist view to a constructivist view will, however, be implicitly explained. I argue, for instance, that what gives us reasons to want certain things are features of the objects of desire. This is something that followers of constructivism would not agree with. It can also be noted that when I argue in favour of a non-reductionist view on normative concepts, I take myself to be answering objections against and making positive claims about the realist version of externalism rather than the constructivist version. There are, despite that, arguments that I use against the internalist view that could be used by a constructivist externalist or some other kind of externalist as well.

<sup>4</sup> There are several ways of classifying reasons that I will not talk about. Nagel speaks of “objective” and “subjective” reasons in *The possibility of altruism* Oxford: Clarendon Press, 1970. Gibbard distinguishes motivating reasons from “potential” reasons in *Wise choices, Apt feelings* Oxford: Clarendon Press, 1990. Scanlon, in his *What we owe to each other* Cambridge: Harvard University Press, 1998, contrasts “objective” and “operative” reasons. And so on. Too many distinctions will only cause confusion.

<sup>5</sup> This example is one that Derek Parfit uses. I have, however, added some further details to it.

<sup>6</sup> In saying this I am not meaning to exclude the possibility of *akrasia*. Surely we can fail to act on what we take to be good reasons to act.

<sup>7</sup> Wlodek Rabinowicz has pointed out to me that while what I claim in my example above (where an explanatory reason isn't a motivating reason) might be true, it is also possible that the reverse of this might be true. He made the following comment: “There are (---) cases in which motivating reasons are not explanatory. Example: If I believe that X needs help, this might provide me with a motivation to help. (If need not, of course. It might also leave me cold.) But if I have this motivation, but help X by pure reflex, then my motivating reason is not an explanatory reason. Or, to take another example, if I do have this motivation to help, but abstain from helping, for example because I also have a stronger motivation against helping X (say, helping him would be too onerous), then my motivation to help does not explain my behaviour. So even when behaviour is intentional, some motivating reasons need not be explanatory.” Surely there can be *some* motivating reasons that aren't

---

explanatory when we act intentionally, but while this may be true, I still believe that other motivating reasons will be explanatory when we act intentionally.

<sup>8</sup> That I have a reason to x is true because some fact obtains. Because I believe that this fact obtains, I believe that I have a reason to do x. I return to this under the heading *The relation between normative and motivating reasons*

<sup>9</sup> Once again: Because I believe that some fact obtains, I believe that I have a reason to act. Because I believe that I have this reason it is likely that I become motivated to do this thing, and because of this it may be the case that I act.

<sup>10</sup> *Practical Reality*, (PR) p. 167 Oxford: Oxford University Press, 2000

<sup>11</sup> I have, up until now, spoken of things being good reasons for acting in order to avoid confusion. Now that I have mentioned this I can drop the addition of ‘good’.

<sup>12</sup> The addition may have rhetoric or stylistic value though.

<sup>13</sup> *IER* P. 101. I write ‘do x’ instead of ‘phi’. Even if Williams points out that *phi* stands for some verb of action I want to make it explicit that it is reasons for actions that he is talking about.

<sup>14</sup> *IROB* p. 35 My italics

<sup>15</sup> p. 40 My italics

<sup>16</sup> However, we should note that externalism about reasons does not, by itself, commit one to the non-reducibility thesis.

<sup>17</sup> Parfit, who I in many ways follow, calls his version of this view ‘Practical Realism’. That I prefer to call it Normative Realism has to do with that what we are discussing here is normative reasons and the nature of normativity. Moreover, calling this view Practical Realism may be misleading in the sense that it may make it appear that this view is only a view about reasons for acting. As I will argue, I believe that the most important reasons are reasons for desiring.

<sup>18</sup> *IROB* p. 35

<sup>19</sup> *IER* pp. 102-104

<sup>20</sup> *IER* p. 102

<sup>21</sup> This could be a non-reductionist internalist view. I will later return to the possibility of such views.

<sup>22</sup> my italics

<sup>23</sup> *Postscript to IER in Varieties of Practical Reasoning* ed. Millgram, Elijah Cambridge: The MIT Press, 2001. p. 91

<sup>24</sup> *IROB* p.36

<sup>25</sup> *Rediscovering Reasons* work in progress

<sup>26</sup> *Values, Reasons and the Theory of Persuasion*, in *Ethics, Rationality and Economic Behaviour* eds. Farina, Francesco, Hahn, Franic and Vannucci, Stefano, Oxford: Clarendon Press, 1996. p. 67

<sup>27</sup> p. 68

<sup>28</sup> p. 68

<sup>29</sup> I return, later on, to what I mean by ‘intrinsically irrational’.

<sup>30</sup> *IER* p. 105-106

<sup>31</sup> *IROB* p.37

<sup>32</sup> pp. 37-38

<sup>33</sup> *IROB* p. 36

<sup>34</sup> P. 36. My italics.

<sup>35</sup> See *Principia Ethica* Cambridge: Cambridge University Press, 1993

<sup>36</sup> Sidgwick suggested this kind of argument when he discusses analytical utilitarianism. Note 1 on p. 26 in *The Methods of Ethics*. Cambridge: Hackett Publishing, 1981

<sup>37</sup> *Rediscovering Reasons*

<sup>38</sup> *Rediscovering Reasons*

<sup>39</sup> I hope to do so elsewhere. Later on, under the heading *What do reasons do and why are there reasons?*, I will, however, make some further brief remarks about what I take normative properties to be like that attempts to show why we should accept a non-naturalist view.

<sup>40</sup> *Rediscovering Reasons*

<sup>41</sup> Shelly Kagan seems to endorse such a view in his *The limits of morality* Oxford: Oxford University Press, 1989. p. 388

<sup>42</sup> Found in Hare’s *Moral thinking* Oxford: Clarendon Press, 1981

<sup>43</sup> Even if the concept of ‘soundness’ perhaps best is understood as evaluative, the appeal to sound deliberation, i.e. the claim that there should be a sound deliberative route, is clearly normative.

<sup>44</sup> *IER* p. 103

<sup>45</sup> P. 39

<sup>46</sup> *IER* p. 106

---

<sup>47</sup> p. 106

<sup>48</sup> I'll discuss this further in the next section.

<sup>49</sup> Internalism proposition (iii) above.

<sup>50</sup> *IER* p. 102

<sup>51</sup> Actually, Dancy does. I will, after this section, discuss what Dancy has to say about this.

<sup>52</sup> My *italics*.

<sup>53</sup> *IER* p. 109

<sup>54</sup> Korsgaard writes: "Practical-reason claims, if they are really to present us with reasons for action, must be capable of motivating rational persons." in *Skepticism about Practical Reason* p. 317 in *Creating the Kingdom of Ends* New York: Cambridge University Press 1995.

<sup>55</sup> Such a theory, I said earlier on, claims that we act because we are in certain mental states.

<sup>56</sup> *PR* p. 167

<sup>57</sup> p. 167

<sup>58</sup> "The difference between false and true belief on the agent's part cannot alter the *form* of the explanation which will be appropriate to his action" *IER* p. 102

<sup>59</sup> Just believing *anything* isn't likely to get you running. The important thing in this case is that he believes that he is being chased by a tiger. This, i.e. the content of his belief, is what he takes to give him a reason to run. Thus we can say that, while he takes himself to be running *because* he is being chased by a tiger, the *explanatory* reason why he is running is that he has a certain belief. It is the fact that he believes something that motivates him to run. The explanation of why he is running is thus that he takes himself to be having a reason to run. So, remembering that explanatory and motivating reasons aren't necessarily the same things will help us see what the relation between motivating and normative reasons is like.

<sup>60</sup> I hope to discuss these things at greater length elsewhere.

<sup>61</sup> Quoted by Parfit in *Rationality and Reasons* in *Exploring Practical Philosophy*, eds. Egonsson, Dan et al Ashgate, 2001 p. 25

<sup>62</sup> P. 25

<sup>63</sup> *Intentions, Plans and Practical Reason* Cambridge: Harvard University Press, 1987

<sup>64</sup> Some non-cognitivists might tell us that there cannot be any normative beliefs. There can, they say, only be normative attitudes because there is no such thing as, for example, moral truths that could make our normative beliefs true or false. I believe that such a view is mistaken. Without arguing for this here I will assume that there really can be normative beliefs. It doesn't even matter, I would say, whether there is such a thing as moral truths, because even in the absence of such truths we could still have normative beliefs.

<sup>65</sup> Or to want, bring about, have a pro-attitude towards this thing. I will not discuss this much further here.

<sup>66</sup> I don't mean by this that these claims have the same meaning. Rather, what I say is that there is a relation of symmetrical implication.

<sup>67</sup> *Reasons* in *Reason and Value: Essays on the Moral Philosophy of Joseph Raz*, eds. R. Jay Wallace, Michael Smith, Samuel Scheffler, and Philip Pettit, Oxford University Press. The version I have of the text is, however, one that may not be the final version. See: <http://users.ox.ac.uk/~sfop0060/pdf/reasons.pdf>

<sup>68</sup> *Rationality and Reasons* p. 27

<sup>69</sup> Gibbard, in *Wise choices, Apt feelings*, endorses a view that has similarities with Parfit's.

<sup>70</sup> In *The Strike of the Demon* (forthcoming)

<sup>71</sup> Though I haven't here included desires that are both intrinsically and instrumentally rational, I believe that there can be such desires as well.

<sup>72</sup> There is actually a need for further qualification here, as this definition of what it is for desires to be intrinsically rational appears to ignore the difference between two kinds of cases. In cases of the first kind it is intrinsic facts about the object of desire that give us reason to have this desire, whereas it, in cases of the second kind is extrinsic facts about the object of desire that give us reasons to have this desire. Since I haven't at this time been able to come up with any good enough names for these separate kinds of rational desires, I here use, well aware of its shortcomings, this wider definition.

<sup>73</sup> Such objections are discussed in *The Strike of the Demon*

<sup>74</sup> A claim like this obviously needs to be explained further. My ambition is to do so in a paper that will also deal with the kind of objections against realism that I've mentioned above.

<sup>75</sup> Some writers, it seems, take normativity to be some motivational force. Korsgaard, for instance, writes that "the normativity of obligation is, among other things, a psychological force". This, I believe, is a mistake. Even if normative reasons ought to motivate us, we should accept that, while facts about our motivating reasons are psychological facts, when we speak of normative reasons, what we are trying to decide is how we *should* be motivated. See Korsgaard's *The Meta-Physical Foundation of Normativity* (John Locke-lectures: lecture 1) <http://www.people.fas.harvard.edu/~korsgaard/Korsgaard.LL1.pdf>

---

<sup>76</sup> For helpful comments on earlier drafts I should like to thank Wlodek Rabinowicz and the participants of the seminars in Practical philosophy at Lund University.