

Parfit's Case against Subjectivism¹

David Sobel

June 23, 2009

DRAFT-Comments most welcome

Derek Parfit, in the early chapters of his magnificent *On What Matters*, argues that all subjective accounts of normative reasons for action are false and should be rejected in favor of some objective account. Parfit develops three main arguments in this part of the book; The Agony Argument, the All or Nothing Argument, and the Incoherence Argument. The Agony Argument and the Incoherence Argument target subjective accounts while the All or Nothing Argument concludes that no desires provide reasons. Today my focus will be on the Agony Argument which I take to be the most fully developed and his most favored of the three. Only of this argument does he say that he thinks it “decisive”. It provides Parfit’s most sustained and novel critique of subjectivism. Elsewhere I have responded to a more fully elaborated version of The Incoherence Argument offered by David Enoch.²

“Who could possibly deny that the nature of agony gives us reasons to want to avoid being in agony, and that the nature of happiness gives us reasons to want to be happy?”, Parfit asks. He claims that the Agony Argument shows that “such claims *must* be denied by those who accept subjective theories about reasons.” [2, 66]

¹ I am very grateful to Janice Dowell, Derek Parfit, Connie Rosati, and an audience at the University of Oxford for helpful feedback on the ideas in this paper.

² I argue that the arguments of David Enoch’s “Why Idealize?,” *Ethics*, 2005 do not cast doubt on the truth of subjective accounts in my “Subjectivism and Idealization,” *Ethics*, 2009. Parfit says in a footnote that his Incoherence Argument presses concerns similar to Enoch’s.

The Agony Argument offers a two premise argument to the conclusion that all subjective accounts are false. The first premise is that we have current reasons to avoid future agony. The second premise is that subjective accounts cannot vindicate this fact. So, the argument concludes, subjective accounts must be rejected in favor of some objective account of normative reasons. I will accept the first premise of this argument and that the argument is valid. The main thesis of this paper is that subjectivists can account for our reasons to get pleasure and avoid agony. I will therefore conclude that the Agony Argument does not justify the rejection of subjective accounts. Independently of that, I will examine Parfit's understanding of the distinction between objective and subjective theories. A secondary thesis of this paper is that Parfit offers a surprisingly narrow understanding of subjectivism such that even if his critique were successful, this would be bad news for fewer theories than we might have thought.

Parfit has a variety of arguments against naturalistic accounts of what it is to be a reason. He argues that such accounts must fail and that what it is to be a reason is an irreducibly normative non-natural property. However, what makes it the case that something is a reason or what provides a reason, he insists, can be a natural property. I am here hoping to isolate Parfit's worries specifically against subjective accounts understood as accounts of what provides reasons. He claims the Agony Argument refutes subjectivism so understood.³ This is the claim I address here.

Parfit sometimes suggests that the subjectivist theories of some prominent philosophers, including Bernard Williams, Richard Brandt, and John Rawls, are best understood as disguised tautologies or definitional truths rather than substantive theses about what we have normative reason to do or what is good for us. I will here always understand subjectivism to be offering a substantive account of an area of discourse. The area of discourse we will be concerned with here will be what has confusingly come to be called objective normative reasons. Objective normative reasons depend on the facts, not one's epistemic access to the facts. Subjective normative reasons depend on one's epistemic situation.

³ Note to the footnote where Parfit says he is addressing such a conception of subjectivism in this part of the book.

The Agony Argument

Parfit's Agony Argument shares an argumentative structure with a more familiar objection to subjective accounts which we might call the "Traditional Extensional Argument". It claims that we have reasons to be moral even when there is nothing that matters to us that would be served by doing so. If this were true it would refute subjectivism since, as I will understand it here, subjectivism entails that the only thing that ultimately provides reasons is a privileged set of what matters to us or our concerns. Mark Schroeder helpfully labels objections of this type the "Too Few Reasons Objection". The objection is that subjective accounts cannot vindicate all the reasons we are confident there are.⁴

Parfit offers a variant on this argumentative scheme which relies on less controversial premises than the Traditional Argument. The Agony Argument maintains that we have reasons to want to get feelings that we like and to avoid agony. We have such reasons even in situations in which we have no current desires which would be served by avoiding agony. Subjectivism is incompatible with the truth of this claim and so must be rejected. An advantage of this argument over the Traditional Extensional Argument is that it is more obvious and more universally accepted that we have reasons to get what we like and to avoid agony than it is that we have reasons to be moral. Many subjectivists have been willing to accept that agents only contingently have reason to act in accord with morality's requirements. It would be harder to claim that future agony only contingently provides reasons to avoid it.

To evaluate this argument we need to understand what Parfit means by agony and pleasure. Parfit accepts that current phenomenological states only count as agony or pleasure if they are liked or disliked. And this liking or disliking itself is not a reaction we have any reason to have. We simply have such

⁴ Mark Schroeder argues that subjectivists have real hope of vindicating the claim that all agents have strong reasons to be moral in his *Slaves of the Passions*, Oxford 2007. I find grounds for resisting this claim in my review of this book posted on 4/25/09 on the *Notre Dame Philosophical Review* <<http://ndpr.nd.edu/review.cfm?id=15905>>

reactions to some phenomenology and not to others. But given that we like or dislike certain phenomenology, he claims that we have reasons to get what we like and to avoid what we dislike.

But subjective accounts cannot vindicate this truth, Parfit claims, for two key reasons which will be the focus of much of this paper. First, Parfit claims, likings are importantly different from desires and so subjectivists cannot appeal to the reason-givingness of likings. Second, subjective accounts can only grant reason-giving authority to desires I currently have or would currently have after informed deliberation. But a person might lack any current desire, even after informed deliberation, which would give weight to future likings or desires. And so subjective accounts must admit that they cannot capture the thought that one's future agony necessarily provides one with reasons now to take steps to avoid it. I will argue that we should be unconvinced by Parfit's arguments for these two claims. Deprived of these two key claims, the Agony Argument fails.

One might resist Parfit's Agony Argument either by denying the claim that likings provide reasons or by claiming that subjectivists can vindicate such reasons. I will adopt the latter strategy. When we understand pleasure and agony as Parfit does, such that they are constituted by favorable and unfavorable responses in cases where the agent is accurately informed about the object of her response, I agree with him that such states provide reasons. But I will claim that Parfit has not given us sufficient reason to doubt that subjectivists can accommodate this fact.

Parfit claims that likings are importantly different from desires. While he acknowledges that likings are in some ways "desire-like" [63] he argues that there are key differences between the two states and that it is mistake to think of liking as a kind of desiring. Parfit relies on the claim that not only are

likings different from desires but also that subjectivists cannot grant authority to likings. Indeed the former claim is useful for the Agony Argument only if it helps establish the latter claim.⁵

Parfit tells us he will understand the category of desire as broader than it typically is understood in English. Parfit writes “The word ‘desire’ often refers to our sensual desires or appetites, or to our being attracted to something, by finding the thought of it appealing. I shall use ‘desire’ in a wider sense, which refers to any state of being motivated, or of wanting something to happen and being to some degree disposed to make it happen, if we can.” [56] Parfit contrasts desires in this sense with

hedonic likings or disliking of certain actual present sensations that make our having these sensations pleasant, painful or in other ways unpleasant.... It is sometimes claimed that these sensations are in themselves good or bad in the sense that their intrinsic qualitative features or what they feel like give us reasons to like or dislike them. But we do not, I believe, have such reasons.... Whether we like, dislike, or are indifferent to these various sensations, we are not responding or failing to respond to any reasons.

When we are in pain, what is bad is not our sensation but our conscious state of having a sensation that we dislike. If we didn’t dislike this sensation, our conscious state would not be bad.

When we are having some sensation that we intensely dislike, most of us also strongly want not to be in this conscious state. Such desires about such conscious states we can call meta-hedonic. Many people fail to distinguish between hedonic likings or disliking and such meta-hedonic states. But these states differ in several ways. [63-4]

Parfit goes on to offer four differences between meta-hedonic desires and likings. First, he says “What we dislike is some sensation. What we want is not to be having a sensation that we dislike. Our desire could be fulfilled either by our ceasing to have this sensation, or by our continuing to have it but ceasing to dislike it. No such claims apply to dislikes, which, unlike desires, cannot be fulfilled or unfulfilled.” [64] Second, we can only like or dislike sensations that we are experiencing. He writes “We cannot now dislike this future pain. ... Unlike our meta-hedonic desires, our hedonic liking or disliking

⁵ In conversation Parfit is inclined to place less weight on the claim that likings are not desires and more weight on the claim that subjectivists cannot appeal to the reason-giving power of future desires. I largely focus on the arguments in the text of May 9, rather than conversations with Parfit.

cannot be aimed at the future, or at what is merely possible. That is another reason why I do not call these mental states desires.” [64] Third, our likings can create reasons or “make some of our conscious states good”. [64] Our desires cannot do so. Fourth, while we have no reasons to have or not to have our likings, we can have reasons to have meta-hedonic desires. For example, Parfit claims, we have reasons to desire not to be in agony.

Recall what is relevant here. Parfit is offering a case that likings are not a kind of desire. Presumably this is because it is assumed that if likings were a kind of desire, subjective accounts would be in a good position to account for their reason-giving force. In making a case that likings are not a kind of desire, it is not helpful to claim that likings create reasons and desires do not for that merely presupposes that which is at issue, namely that the two states are fundamentally different without arguing for that conclusion. So the third difference Parfit points to begs the question in this context. If it is true that likings can and desires cannot provide reasons, this must be due to something that is different about the two states. Parfit is supposed to be explaining what that difference is.

Parfit tells us that there are two ways of satisfying a meta-hedonic desire, either by no longer minding a sensation you are feeling or stopping to feel that sensation. Likings are claimed to have no comparable feature. This feature of meta-hedonic desires is not a feature of all desires. Some desires cannot be similarly satisfied in two ways. If I want to smell a flower, I cannot satisfy this desire either by smelling the flower or by losing such a desire.

Parfit also says that desires can be satisfied but likings cannot. But both states are favorable psychological attitudes that have objects which we can either get or fail to get. Perhaps to Parfit “liking” suggests a relatively persisting favoring of a state which cannot be fulfilled by any particular instance of getting what one likes whereas desire seems as though it is for a particular occasion. But this is a pretty shallow difference between desires and likings. A person who has a settled and permanent preference for Guinness over ale would not, in this sense, satisfy such a preference by getting one. Or perhaps the

thought here is that likings do not take propositional attitudes whereas desires and the like do. While I admit that this is plausible, I am not sure it is true. We do not, after all, really literally like chocolate ice cream. Rather we like the taste of it. And I do not, in the relevant sense, like your tasting it but rather what I like is the sensation I get when I taste it. I am unsure what is at stake in insisting that it mischaracterizes things, rather than just sounds awkward in English, to say that what I like is that I am experiencing a certain flavor of sensation. However, even if this is a difference between likings and desiring, it does not help us see why the one could provide reasons and the other not. Further, as we will presently see, the key claim here is not that likings are a kind of desire but rather that the two states are strongly correlated. And this could be true even if likings are not a kind of desire.

Parfit says that we can only like current sensations and thus likings are unlike meta-hedonic desires which can aim at future states. But it seems that some subset of desires could share this feature. Perhaps desires which are directly satisfied only because of the phenomenology one is currently experiencing, for example, are similar to likings in this respect.⁶

But most importantly, there are many different kinds of desires. Our question is not whether likings are different from some kinds of desires. Surely they are. Our question is whether likings are a sub-category of desire. Likings and meta-hedonic desires could be different, yet both sub-classes of desire. Most of the differences that Parfit points us to aim to show that there are differences between likings and meta-hedonic desires. But we can grant this and still claim that likings are a kind of desire. To be on point, we would need to be given an argument for thinking that if likings are not meta-hedonic desires, they could not be some other kind of desire. But we are given no such argument. I will shortly consider reasons to doubt this claim. But for now notice that we have exhausted the differences between likings and desires that Parfit points to, yet seen no reason to think that likings are not a kind of desire, want, or preference.

⁶ This proposal will be sharpened later in the paper.

If we thought that likings were a kind of desire, we would be unlikely to think that they were meta-hedonic desires. The object of a liking is a particular kind of feeling; say the taste of chocolate ice cream. But the object of meta-hedonic desires is not a particular phenomenology but any phenomenology which is favored in a certain way. If likings were a sub-class of desires, meta-hedonic desires would take that class of desires as its object. So anyone claiming that likings are a kind of desire should allow that likings are different from meta-hedonic desires.

There are some reasons to think that likings are a kind of desire or preference. Presumably if liking something provides a reason, liking it more provides a greater reason. This seems to be the thinking behind Parfit finding it obvious that we have a significant reason to avoid agony. I understand agony to be a sensation that one very strongly dislikes. So to understand the strength of reasons that liking provides we will have to compare the strength of liking. How might this be done? Any sensible way of ranking the degree of pleasantness of the options would seem to require that we put the different options in some way before the agent and look to the agent to render a verdict about the comparison. This suggests that perhaps we should mimic a popular story told by subjectivists and say that one should construct an idealized version of oneself who has had, and retained an accurate impression of, different things that one likes and dislikes. Then we might say that X is liked more than Y iff an agent intrinsically prefers experiencing state X over state Y where the preference is for intrinsic qualitative features of the phenomenological state.

Perhaps it will be said that liking more is distinct from the sort of preference just mentioned. I doubt this, but suppose it is so. Still it is not clear what it could mean to say that one likes the feel of phenomenological state x more than y but does not have the above sort of preference for x over y while fully and accurately informed about what the phenomenology is like. Since the structure of Parfit's complaint is extensional in form, it actually does not matter for present purposes if liking more is identical with the above sort of preference or if the two states merely necessary co-vary. So to resist what I am

urging here one would have to claim that liking more can be captured in some way that does not co-vary with the above sort of preference.

Roger Crisp offers the following example alleged to show exactly this. “Imagine that I have never experienced serious pain. I might, during my first experience of it, desire it for its novelty, at least for a short time—and there is no need to think that I must somehow be enjoying the novelty. I desire the pain for how it feels, but there is no enjoyment here.”⁷ But I think this example is not telling. The person in his example desires the phenomenology because it is novel to them, not for an intrinsic qualitative feature of the phenomenology.

Let us step back a bit from the argument for and against the claim that liking is a kind of desire. This has seemed an important question for us to settle largely because we have so far been making two assumptions; first that if likings were a kind of desire, that would ensure that subjectivism could explain their reason-giving power, and second that if likings are not a kind of desire they cannot. But both of these assumptions seem false. Here I focus on the first assumption. I will address the second in the next section.

Subjectivists have wisely tended to deny that every desire provides a reason, claiming rather that some kinds of desires provide reasons and others do not. Thus a person might allow that likings are a kind of desire, yet claim that they are not the sort of desires that subjectivists can plausibly claim provide reasons. If this were true then granting that likings are a kind of desire would not help the prospects for subjectivism.

⁷ Crisp, *Reasons and the Good*, Oxford University Press, 2006, p. 107. In conversation Crisp suggests that perhaps the novelty of the sensation is part of the sensation. While I agree that sensations often feel different when they novel, I do not think that there is some common sensation that novel sensations have in common. I am not sure Crisp is disagreeing with that claim.

As I have argued elsewhere, subjectivists should claim that intrinsic favoring and disfavoring attitudes ground reasons when they are accurately informed about what their object is like. Such desires are more fully for their object as it really is rather than for the object as it is falsely believed to be. If this story were granted, there would be a fairly natural subjectivist story about why, if likings were a kind of desire, they would be the sort subjectivists can most plausibly claim have reason-giving authority. Desires for current phenomenology are uniquely accurately and fully informed about their object. Indeed, it might now seem a mark in the subjectivist's favor that we think that the cases in which we have the most confidence that our desires carry normative authority are cases in which we are most confident we have excellent access to accurate information about what certain options are like. Matters of mere taste (where desires are commonly allowed to carry authority) tend to be cases where we have uncommon access to the relevantly informed vantage point. If likings were a kind of desire, subjectivist could account in a natural way for the reason-giving power of such states in a way that fits well with their broader approach. In other words, if likings were desires they would be just the sort of desires that subjectivists can most plausibly grant authority to; namely those desires which are accurately informed about their object.

The Status of Future Desires

Parfit has a second argument for the key second premise of the Agony Argument that subjectivists cannot account for our current reasons to avoid future agony. Parfit maintains that subjective accounts can only claim that an agent has a reason now to avoid future agony if the agent currently has, or would currently have after ideal procedural deliberation, a desire to do so. And he thinks it must be left a contingent matter whether or not one's future agony would generate such a current desire.

He writes

I know that some future event would cause me to have some period of agony. Even after ideal deliberation, I have no desire to avoid this agony....

Since I have no such desires or aims, subjective theories imply that I have no reason to want to avoid this agony, and no reason to try to avoid it, if I can....

This case might be claimed to be impossible, because my state of mind would not be agony unless I had a strong desire not to be in this state. But this objection overlooks the difference between our attitudes to present and future agony. Though I know that, when I am later in agony, I shall have a strong desire not to be in this state, I might have no desire now to avoid this future agony....

It might next be claimed that my predictable future desire not to be in agony gives me a desire-based reason now to want to avoid this agony. But this claim cannot be made by those who accept subjective theories of the kind that we are considering. These people do not claim, and given their other assumptions could not claim, that our *future* desires give us reasons. [3, 78]

Parfit's claim that subjectivists can appeal only to current desires is not argued for but rather functions as a stipulation. Later we will consider what might be motivating this stipulation, but for now, let us accept it.

One reason Parfit thinks subjectivists have trouble accounting for our current reason to avoid future agony is that what makes a state agony is not a desire. We considered his case for this earlier. But note that if there is to be a distinct, second argument for this claim, that second argument must not turn on the claim that future agony is not a kind of desire. That would just be another application of the thought that likings are not desires. And in that case there would be no need to talk about future agony. Parfit could make his point using current agony.

The new problem posed by the case of future agony concerns the transfer of reasons one will have in the future to reasons one has now on a subjectivist account. So Parfit should and does suppose that the subjectivist can account for one's reasons to take one's hand out of the fire when one is actually in agony. Parfit does not dispute that when we are in agony we necessarily have a desire to get out of it and that subjectivists can account for the reason we have to avoid current agony. The new problem posed by the case of future agony is about the insecurity, on the subjectivist account, of the transfer of reasons one will have in the future to avoid agony to reasons one has now to take steps to avoid it in the future.

Why should we think this transfer is insecure on a subjectivist account? As Parfit points out, it is clearly psychologically possible that the belief that we will have a strong desire in the future, say to avoid

hell, fails to produce a strong desire in the agent now to take steps to avoid that situation. In some cases this might be because our current values are hostile to the values we will have in the future. If I am not now vicious but know I will become vicious in the future I may be unmoved by the fact that I will have strong future desires to torture. These cases are difficult. Let's consider only cases where the agent's current values do not speak against the agent's future values. Even in such cases, as the example of Hell shows, we can as a psychological matter fail to be moved by the thought that something will matter to us in the future in a way that is clearly intertemporally irrational. Parfit is clearly right that descriptive psychology will not ensure the proper transfer of desires in cases where we are certain there is a transfer of reasons. In the actual cases we are familiar with, the problem is usually that the future pain is, as Sidgwick put it, "foreseen but not fore felt". The subjectivist suggestion that we provide agents with an accurate and retained impression of what the future agony will be like would surely go a long way to curing at least most actual cases of such irrationality. But Sidgwick thought that this would not solve all possible cases of such irrationality.

If Sidgwick is right, subjectivists cannot rely on the causal impact of accurate information about possible futures to ensure that agents are moved by their future concerns. I will grant this. Given that, the subjectivist can successfully respond to Parfit's challenge only by building in transfer principles into their account of ideal procedural deliberation. Parfit argues that, given their commitments, subjectivists cannot do this. He argues that "Subjectivists cannot claim that, if we were procedurally rational, we would want to avoid all future agony." [3, p. 83] Parfit considers arguments to the contrary by Michael Smith. (We will see later that Smith counts as a subjectivist for present purposes). Smith argues that those who are future Tuesday indifferent, for example, are making arbitrary distinctions and that the desires of such a person are less coherent and unified than the desires of someone who is not future Tuesday indifferent. Smith is thinking that the subjectivist can help themselves to the claim that ideal procedural deliberation will result in desires that are coherent and unified. If all that were right, subjectivists could say that a person who does not give weight to future agony, but only to present agony, is making an arbitrary

distinction. Such a person's preferences are less coherent and unified and it is for these reasons that an ideal procedural deliberator would not make such arbitrary distinctions and would want to avoid all future agony.

Parfit replies that such arguments are unavailable to the subjectivist. Some of his arguments are specifically against Smith. And I may be persuaded by some of those arguments. However, only one of Parfit's arguments against Smith offers real prospects of generalizing against all subjectivist attempts to claim that ideal procedural deliberation would require one to want to avoid all future agony. Parfit argues that

Our preferences draw arbitrary distinctions when, and because, what we prefer is in no way preferable. It is arbitrary to prefer one of two things if there are no facts about these things that give us any reason to have this preference.... To explain why this preference is arbitrary, we must claim that (1) if some ordeal would be on a future Tuesday, this fact does not give us any reason to care about it less....[M]ost of us would always prefer to have one of two ordeals, if, and because this ordeal would be less painful. To explain why this preference is not arbitrary, we must claim that (2) if some ordeal would be less painful, this fact does not give us a reason to care about it less. (1) and (2) are claims about object-given reasons. Since Subjectivists deny that we have such reasons, these people cannot appeal to such claims, or to the 'minimal principle' that Smith states... [3, 82]

In response to this argument I will first make a case that as formulated the argument is question-begging and so needs reformulation. I offer a reformulation and then go on to offer grounds for resisting what I take to be the general thought Parfit urges here.

Parfit defines objective theories as those that claim that "there are certain facts that give us reasons both to have certain desires and aims, and to do whatever might achieve these aims. These reasons are given by facts about the objects of these desires or aims...." Hence he calls such reasons "object-given" reasons. [chapter 2, 57] He defines subjective theories as those that claim that "our reasons for acting are all provided by, or depend upon, certain facts about what would fulfill or achieve our present desires or aims.... Since these are all facts about us, we can call these reasons subject-given."

[ibid.] We have no reason to have the desires which (allegedly) provide subject-given reasons (ignoring, as I do throughout, state-based reasons).

Parfit's understanding of the subjective/objective distinction is in terms of whether or not there are standards for what one has reason to desire. To the extent that there are, to that extent objectivism is correct, he claims. But, on this way of marking the divide, it is irrelevant what provides those standards. A desire-based reason for P to O would be subject-given at the time the agent has the initial desire to O because she would have no reason to have that desire. But if a later desire to get out of agony provides P a current reason to want to avoid agony in the future, this earlier reason should count for Parfit as an object-given reason. One would have a reason to have this desire. It would then be due to facts about the object of the earlier desire, namely that it will be desired later, which provide the reason to desire it now. To assume that only objectivists can grant authority to object-given reasons is therefore to assume that ideal procedural deliberation cannot require one to care now about the fact that one will later care about something. And this is directly question-begging in the context of the argument Parfit is giving. Thus we must not presuppose, as Parfit does, that merely because a reason is object-based in his sense that only objectivists can grant authority to it. To do so would be to beg the key question about whether subjectivists can claim that ideal procedural deliberation requires us to care now about what we will later care about. If subjectivists can claim this, then they can appeal to object-given reasons in Parfit's sense.

So the problem with the subjectivist appealing to Principle (1) is not that this would be an object-given reason. And, since Parfit needs to grant that subjectivists can account for our reasons to avoid current agony, the principle that the subjectivist really needs to rely on is not (1) but rather something more like what I will call the Reasons Transfer Principle. It states that: If one will later have a reason to

get 0, then one now has a reason to facilitate the later getting of 0. Parfit needs to make a case that subjectivists cannot appeal to a principle such as the Reasons Transfer Principle.⁸

I am forced to speculate. I would guess Parfit wants to say that such principles, according to subjectivism, must themselves be supported and approved by an agent's contingent concerns. But I do not think this is so. Analogously, Christine Korsgaard claimed that subjectivists need a separate principle enjoining one to take the means to one's authoritative ends. She thought that once the subjectivist champions categorical imperatives of the form "take the means to your ends whether you feel like it or not", the subjectivist has lost principled grounds for resisting further categorical imperatives. I think this is mistaken in two ways. First, I don't think saying that one has a reason to take the means to achieve something is an additional claim over and above the claim that one has a reason to achieve something. The former claim seems partially constitutive of the claim that one has a reason. Second, suppose the claim about taking the means was an additional principle. Still, its connection to subjectivism is clear. The reason to serve that principle is that in doing so one will serve one's concerns. The person who thinks contingent concerns provide all ultimate reasons has a principled rationale for letting in claims such as the Reasons Transfer Principle and resisting claims such as that everyone has a reason to aid those in distress. The former serves the agent's concerns whatever those concerns may be. The latter need not.

Sometimes I think people mistakenly assume that a proper subjectivist view can never tell people to do anything that they do not currently want to do. But this is just wrong. Most obviously, the subjectivist is not saying that the principle that one's concerns of the right sort provide reasons itself is up for assessment by anyone's concerns. Consider the actual agent who has not deliberated soundly. She may not want to put money away for retirement. Despite that, subjectivism still claims she has reason to

⁸ The Agony Argument needs to presuppose something like the Reasons Transfer Principle to vindicate the first premise in its two premise argument. There may well be technical problems with my precise formulation. Feel free to substitute a better version sufficient for the purposes of the Agony Argument and I believe my arguments will still apply.

do so. The actual agent may not care about what she would want after ideal procedural deliberation. Still the view says that her reactions after such deliberation determine her reasons. What counts as ideal procedural deliberation is not itself responsive to one's wants or concerns.⁹ That is why subjectivists tend to offer a one-size fits all account of what is involved in ideal procedural deliberation.

I will now offer two reasons to think subjectivists can appeal to principles such as the Reasons Transfer Principle. Additionally, I will argue that even if they cannot, the subjectivist can still get almost everything she wants.

First, the Agony Argument presupposes that the future agent who would feel the agony is the same agent who is currently deciding what to do. It will be her tomorrow who very much wants to get her arm out of the fire.¹⁰ That future concern should now matter to her, according to subjectivism, because it will be a concern of hers. But to help oneself to this metaphysical assumption about identity is not to help oneself to any objectivist principle. The subjectivist has always claimed that it is my concerns of the right sort that provide me with reasons. The concerns I will have tomorrow are my concerns.

Second, the subjectivist who claims that ideal procedural deliberation involves caring about one's future cares is not appealing to what, in the first instance is worth caring about. This is a procedural rather than substantive requirement. The requirement is not responsive to the content of one's future concerns and whether they are worthy of being desired, but rather only that one will come to care about it in the right way. The principle that one should now care about what one will later care about gives one no guidance until one starts to care about this rather than that for no good reason. It is in this sense that I am saying that such a view borrows no objectivist principles about what is worth caring about in the first instance. This claim involves only the thought that if one will care about something later, one should now

⁹ But see Rosati.

¹⁰ It may be that in cases of brain division or branching that our assumption that it is my concerns that provide my reasons become problematic, and so requires adjustment. But if so, this will have nothing to do with the Agony Argument.

care about that fact. This seems continuous with the idea that one's passions set the ultimate goals and further reasons are hostage to what promotes our ultimate goals. Reason is still the scout or slave to the passions. Can it really be said that it is a distinctively objectivist principle that one should act so as to maximally comply with one's subjectively determined reasons over one's life?

Parfit tells us a bit about how to mark the boundaries between ideal procedural deliberation and a more substantive conception. He tells us that we "can be procedurally rational whatever else we care about" [3, p. 81] and, seconding Rawls, he adds that "knowing that people are rational, we do not know the ends they will pursue, only that they will pursue them intelligently."¹¹ If there were no standards for what we should ultimately want in the first instance but only standards for how to intelligently achieve what one wants over time, that seems to be in accord with these platitudes about ideal procedural deliberation.¹² The view I am suggesting claims that there are procedural principles that determine one's authoritative ends but these principles do not require us to have any particular substantive end. Further there are principles that require us to effectively pursue our authoritative ends over time, but these principles require us to effectively pursue whatever we will want in the relevant way without importing any objectivist principles about which ultimate ends are worth wanting.

It may be that Parfit is thinking that to be properly formal or procedural in the relevant sense, a requirement must not place constraints on acceptable outputs of deliberation but only provide inputs to deliberation. So, for example, one might think that the requirement that one's desire be informed in procedural in this sense because it does not insist that only certain substantive ends can be desires in the

¹¹ Parfit, [3, p. 81] citing Rawls 1996, p. 49

¹² Note that the problem Parfit is pointing to should not be understood to merely be lack of interpersonal coherence in reasons claims or collective suboptimal satisfaction of reasons overall. For if these were seen as the problems, they could be fixed by saying that the relevantly authoritative desires are for whole life paths rather than for options within a life. The existence of organic unities already recommends this path. Rawls' subjectivist proposal took this form as have some subjectivist-friendly proposals in decision theory.

relevant way once one is so informed. However, Parfit might say, the requirement that one cares now about one's future cares is a requirement about the acceptable output of deliberation and so not procedural.¹³ In response, I would say that surely the requirement that one's desires be, for example, transitive has always been taken to be a formal or procedural requirement fully compatible with a subjectivist approach. Yet this is a requirement on acceptable outputs of deliberation. Thus I think the way I carve up the procedural/non-procedural divide above is more intuitive and fits better with standard usage.

Finally, suppose I am wrong, and subjectivists must abandon subjectivism to make claims like the Reasons Transfer Principle. The subjectivist would then be forced to abandon the label "subjectivism", but look at what they get in return. The resulting objectivist view would claim that all our reasons ultimately derive from concerns we have no reasons to have and that additional derived reasons are restricted to that which provides clever ways to achieve lives that involve getting as much of what we really want over time as possible. Parfit is here making no case that such a view would be false, but rather only that it would not count as subjectivist. In that case, I urge subjectivists to get over the loss of the label and accept such a view.

Parfit's Account of the Objective/Subjective Distinction

That concludes my direct reply to the Agony Argument. Now I want to consider how Parfit understands the distinction between objective and subjective theories and show that he has a surprisingly narrow understanding of subjectivism.¹⁴ If I am right, even if the Agony Argument were correct, this would be bad news for fewer theories than we might have suspected.

¹³ Parfit suggested this line in conversation.

¹⁴ In truth, I think Parfit's understanding of "subjectivism" is in some ways surprisingly narrow and in other ways surprisingly broad. I do not mean to suggest that Parfit's understanding of subjectivism is so narrow as to be of

Parfit tells us there are two main kinds of views about practical reasons, objective and subjective. Recall that he tells us that objective views claim that there are facts about the objects of our desires and aims which provide reasons to have these desires and aims. Subjectivists, on the other hand, claim that our reasons are provided by “certain facts about what would fulfill or achieve our present desires or aims.” [2, p. 57]

But what explains why desires, aims (and, Parfit sometimes adds, intentions) are the only attitudes on this list of attitudes that the subjectivist can champion as reason-giving? Earlier I argued that likings are plausibly a kind of desire. Now I am saying, even if we assume that likings are not a kind of desire, it is still an interesting question whether likings belong on the list of attitudes which naturally should be classed as subjectivist-friendly.

So how should we approach the question of whether likings should be classed among the subjectivist-friendly set of attitudes? Parfit does not address this question. I think it useful to begin with an attempt to interpret what subjectivists and objectivists have been disagreeing about. Here is an interpretation of this dispute which is simple, clear, and seems to me at least close to what people have meant to be arguing about.

The central question to which subjectivists and objectivists have been giving competing answers is: What is the ultimate source of our reasons? Is it merely that we happen to favor certain options and disfavor others or are there standards independent of such pro and con reactions which determine what we have reason to do? Do our favoring and disfavoring attitudes guild and stain the world with reason-providing status or do our options already have the reason-providing status independently of our happening to go for some options and not for others? I think this is the best way to understand what the dispute between subjectivists and objectivists has been about. But if that is right, what distinguishes them

little importance. As we will see, I think he means by “subjectivism” what I mean by Williamsian Internalism. Williamsian Internalism is a widely held and quite influential view.

is that the former regards our pro and con attitudes as the ultimate source of our reasons. And then we would naturally class both liking and desiring as in this class of attitudes. Thus, if this account of the distinction between objectivism and subjectivism gets at the heart of the debate, then it does not really matter whether likings are a kind of desire. All that matters is that liking and desiring both count as favorable reactions to options, and so those who grant that such states are an important source of reasons, such as Parfit, will be committed to a serious subjectivist component in their account of what provides reasons. Call subjectivism so understood “pro attitude” subjectivism.

Of course, this is not how Parfit understands the distinction between subjectivism and objectivism. He restricts the subjectivist to the current attitudes of desire, intention, and aims. He does not offer a general intuitive understanding of what the subjective/objective dispute is about such that we might in a principled way think about the question of whether likings, assuming that they are not a kind of desire, belong on the list of subjectivist-friendly attitudes. However, we can find clues in other parts of *On What Matters* as well as in his paper “Normativity”. There Parfit aims to show us that a key mistake of 20th century discussions of normativity is that many have “conflated normativity and motivating force”¹⁵ [337] I think it illuminating to understand Parfit as thinking that subjectivists are the people who have made exactly this mistake. And certainly many prominent thinkers in the Humean and Kantian traditions have wanted to cash out normativity in terms of motivational states. It should be noted that Parfit finds Kantians just as guilty of this conflation as Humeans.

¹⁵ See chapter 24, p. 507 Parfit considers four conceptions of normativity. He writes “On a third conception, normativity involves actual or possible motivation....We ought, I shall argue, to reject this motivational conception.” Most of Parfit’s discussion of this motivational conception is aimed against such a conception of what it is to be a reason. I speculate that Parfit thinks such a conception just as confused when offered as an account of what provides reasons. Notice that Parfit is not always careful in his formulation of what the key mistake is. At one point in discussing Railton, Parfit complains that in Railton’s work “The concepts normative, recommend, and matter are here conflated with, or reduced to, psychological appeal.” [“Normativity.” 338]. But surely likings are a kind of psychological appeal. For Parfit’s case against desires to not also tar likings, he must limit his complaint to confusing the normative with motivational states, rather than all species of psychological appeal.

I think Parfit stipulatively understands subjectivism to include only views that limit their list of normatively authoritative attitudes to motivational attitudes. This, for him, is what unites the list of attitudes he offers the subjectivist; desires, intentions, and aims. And this is why it is important to him that likings are not a kind of desire. He is thinking that likings are not essentially tied to motivation. And Parfit's goal is to show how we can capture many of the reasons that seemed to most obviously require motivationally charged desires, with non-motivational likings. For him, the fight between objectivists and subjectivists is the fight between those who think that "there seems nothing for value to be, on deepest reflection, wholly apart from what moves, or could move, valuers, agents for whom something can matter" and those that dispute this.¹⁶ Call subjectivism so understood "motivational" subjectivism.

Parfit's Agony Argument is attempting to make a case against motivational subjectivists but doing so in a manner which fits quite nicely with pro-attitude subjectivism. The Agony Argument may or may not be a good reason to reject motivational subjectivism but it could not provide a reason to reject pro-attitude subjectivism.¹⁷

We know that Parfit thinks the Agony Argument rules out all subjectivist views and forces us to choose some objectivist view. So given how Parfit defines objectivist views, he would have us call pro-attitude subjectivism a version of objectivism. What we call the view is not important. For the sake of neutrality I will call the view under discussion "pro-attitude psychologism". Pro-attitude psychologism maintains that all of one's reasons ultimately stem from one's favorable reactions that one has no reason to have.

¹⁶ Darwall, Gibbard, Railton, "Toward Fin de Siècle Ethics: Some Trends," p. 176-7. On my reading of him, Parfit must be thinking the word "moves" in the above passage must be understood literally—as a kind of motivation—rather than just saying that the thing in some way must matter to the agent. Surely liking something is a way of that thing mattering to us.

¹⁷ This for two reasons. First, and most obviously, this argument avoids the sting of the claim that likings are not desires. But secondly, such a position could not be committed to what I will call Williamsian Internalism for it claims that non-motivational states can provide reasons.

There is some reason to think that perhaps Parfit missed the fact that his Agony Argument does not rule out pro-attitude psychologism. Parfit thinks that it is no accident that most naturalists are subjectivists about reasons and that naturalists “have strong apparent reasons to accept a motivational account of normativity.” [chapter 26, p. 562] An attraction of subjectivism is its potential to offer a unifying understanding of how a wide variety of kinds of options can provide us with reasons. Parfit tells us that “If Naturalists are not Subjectivists, there is no similar way in which they could explain how such a great variety of facts can give us reasons.” [26, p. 563] But pro-attitude psychologism is surely a naturalist view. And seemingly it has the potential to offer a unified account of what it is about a great variety of options which provides us with reasons. So Parfit needs pro-attitude psychologism to count as a subjectivist view for the purposes of this argument. But he needs it to not count as a subjectivist view so that his Agony Argument could have some hope of plausibly concluding that it had refuted all subjectivist views.

My hypothesis is that because Parfit was focused on motivational subjectivism he did not pay sufficient attention to pro-attitude psychologism. But the existence of pro-attitude psychologism shows us that the only plausible conclusion of the Agony Argument is less exciting than it seemed. The Agony Argument does not rule out the view that all of our reasons are provided by our just happening, for no good reason, to find favor with some options and disfavor with others. If we adopt Parfit’s taxonomy, we can say that subjectivism is refuted. But to my ear that would exaggerate and mislead about what the Agony Argument has accomplished. To vindicate the claim that not all of our reasons are provided by our happening to favor certain options, Parfit will have to go beyond championing reasons that flow from our liking current phenomenological states. He will have to offer us something more like the Traditional Extensional Argument against subjectivism which claimed that we have reasons to be moral regardless of whether doing so answers to anything that happens to matter to us.

Parfit, recall, introduces subjective theories as those theories that only appeal to the authority of present desires or aims. This claim functions as a stipulation. Our assumption that Parfit is only arguing against motivational subjectivism provided an answer as to why he limited the subjectivist to desires, intentions, and aims. But why does Parfit stipulate that only current motivational attitudes can provide current reasons on a subjectivist account? There is an obvious answer which again narrows his understanding of subjectivism.

Williams' influentially claimed that if someone has a reason to ϕ , it must be possible for that person to be motivated to ϕ via sound deliberation. Williams claims that we must maintain an in principle connection between our reasons and what could explain our action, such that if we have a reason to ϕ we might, at least in principle, ϕ because of that very reason. It is taken for granted that when we intentionally ϕ that needs to be explained by our desires. To keep alive this connection, he thinks we must understand our reasons to be that which would motivate us, at least after ideal procedural deliberation. Thus Williams thought that if you would not have a motivation to ϕ after sound deliberation, then you do not have a reason to ϕ . Call that thought Williamsian Internalism.¹⁸

If Parfit has a Williamsian Internalist picture in mind, his stipulation that subjectivists can help themselves only to the authority of current motivational states makes perfect sense. Williams ties reasons to motivations in order to explain how we could act for our reasons. The key thought behind Williamsian Internalism is that normative reasons must be capable of explaining actions, at least after sound deliberation. If we agree with this, we will have to tie current reasons to motivations at the time of the possible action. Future desires could not explain earlier actions unless they generate current desires. So only current desires (perhaps after ideal deliberation) are fit to play the explanatory role that Williamsian Internalism demands.

¹⁸ I myself would not mind just calling this Internalism. I instead call it Williamsian Internalism to highlight that I mean nothing more than what I explicitly attribute to the view in the text above. I do not mean to privilege a Humean rather than Kantian reading of Williamsian Internalism.

I am suggesting that we can best understand Parfit's unexplained stipulation that subjectivists can only appeal to current desires if he is taking subjectivism to just be Williamsian Internalism. This would also explain why the distinction between Humeans and Kantians does not much interest him as both tend to accept Williamsian Internalism. I conclude that understanding Parfit to have in mind Williamsian Internalism when he uses the word "subjectivism" makes sense of an otherwise inexplicable feature of Parfit's view: namely restricting the subjectivist to current desires the agent has after sound deliberation.¹⁹

It is, I hypothesize, because Parfit really has Williamsian Internalism in mind that he is unconcerned with pro-attitude psychologism or with "a temporally neutral desire-based theory" which does not appeal only to the reason-giving power of current desires. [3, p. 67] Of the latter, Parfit merely says "These imagined theories are also very different from the subjective theories we are now considering." [3, 67] This despite the fact that on such views only desires ultimately provide reasons. The Agony Argument, we are forced to conclude, does not aspire to argue against the view that only desires we have no reason to have ultimately provide reasons. Rather, the real target is Williamsian Internalism.

Parfit's stipulation categorizes a surprising range of views that seemed subjectivist, such as the temporally neutral desire-based theory above, as objectivist. Further, if Parfit is equating Williamsian Internalism with subjectivism, then his understanding of subjectivism is idiosyncratic in that it includes many Kantian views. He certainly does target many Kantians as well as Humeans with his Agony Argument. His understanding of subjectivism would be further idiosyncratic in that most think Williamsian Internalism an argument for subjectivism, not the thing itself.

Most would grant, I think, that Parfit's categorization excludes some subjectivist-seeming theories from counting as subjectivist. However, because I think the form of subjectivism that Parfit and I

¹⁹ In an earlier version of *On What Matters* Parfit argued for the claim that subjectivists were committed to appealing only to current desires by attributing to them Williamsian Internalism as I describe it above. Such arguments have now been removed and the restriction now functions as a stipulation.

agree is the most plausible must reject Williamsian Internalism, I think Parfit's understanding of subjectivism to be even narrower than it appears.

The idealized soundly deliberating agent is quite different from the actual agent whose reasons we are concerned with. Maybe the actual agent has a reason to become better informed about some matter but the idealized agent has no such reason. This has led the state of the art subjectivist views to move to what are called "Ideal Advisor" account in which it is the wants of the idealized agent for the actual agent which are the relevant attitudes. But on such a picture, no one need be motivated to do what the theory claims the actual agent has a reason to do. The idealized agent would have no motivation to herself go and read the information she already knows. And the actual agent, of course, may well not be motivated to O as she has not yet deliberated soundly. Thus there is no one on this picture who need be motivated to O. Such a view claims that normative reasons to O may well fail to motivate anyone to O, even the agent who has soundly deliberated. Thus the most promising path for subjectivists to take, namely an idealized advisor account, seems incompatible with Williams' requirement that a reason to O must be capable of motivating an agent to O after sound deliberation. The Ideal Advisor version of subjectivism does not ensure that reasons must be capable of serving as explanations. In the story I just told, no one was motivated to get more information, yet this does not cast down on the thought that if the idealized agent would intrinsically want the actual agent to intrinsically want to O, the actual agent has a reason to O.²⁰ Thus I claim the best versions of subjectivism must leave Williamsian Internalism behind. So the best subjectivist view, the Idealized Advisor account, the view Parfit seems to allow is the most plausible version of subjectivism, is actually not an instance of subjectivism as Parfit understands it.

Conclusion

²⁰ See Robert Johnson's "Internal Reasons and the Conditional Fallacy," *The Philosophical Quarterly* 49, no. 194, Jan. 1999: 53-71. See also my "Explanation, Internalism, and Reasons for Action" *Social Philosophy and Policy* 18:2 (Summer 2001): 218-35 (this paper also appears in *Moral Epistemology*, Cambridge University Press, 2001.)

I conclude that the Agony Argument should not persuade us that subjective accounts are false.

[More?]