

Why Prefer a System of Desires?

JASON A. HILLS

IN his paper “Internal Reasons,” Michael Smith argues that the internalism requirement on a theory of reasons involves what a fully rational version of an agent would desire that his real-world counterpart do. Smith agrees with most aspects of the account of internal reasons given by Bernard Williams in “Internal and External Reasons,” but expands Williams’ account of deliberation to include making our desires *systematically justifiable*. With this addition to deliberation, he argues—contra Williams—that the internal reasons view is non-relativistic in the demands it makes on agents. I will argue, however, that making our desires systematically justifiable is not a necessary component of rational deliberation.

Williams on Internal Reasons

A *reason statement* is a statement about what an agent should do in a particular circumstance. An *internal reason statement* is a reason statement which holds only if the agent has some motive or desire which will be served by his acting in that particular way (Williams 101). Using this definition, Williams argues that internal reasons must relate to the contents of an agent’s “subjective motivational set” (102), sometimes abbreviated as simply “S.” Williams explains that the subjective motivational set of an agent is not limited to what we ordinarily term “desires,” but “can contain such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects” as well (105).

Williams proposes a very basic, “sub-Humean” model for how internal reasons might work: for some agent A, “A has a reason to ϕ iff A has some desire the satisfaction of which will be served by his ϕ -ing” (101). To this

Jason A. Hills is a senior studying philosophy at Brigham Young University. His interests include philosophy of language, logic, and philosophy of mathematics. This essay won first place in the 2008 David H. Yarn Essay Contest.

simple formulation, Williams offers four refinements. As Smith is primarily concerned with the last refinement, I will only briefly mention the others. Williams' first refinement actually comes from the sub-Humean model: "An internal reason statement is falsified by the absence of some appropriate element from [an agent's subjective motivational set]" (102).

Williams makes a second refinement as he considers false beliefs. It is possible for an agent to have desires based upon false beliefs, or to make mistakes about the means necessary to satisfy a desire. For example, an agent might see a bottle of gasoline and think it is a bottle of gin. Perhaps the agent is thirsty and would like to drink the contents of the bottle (or thinks drinking the contents of the bottle will satisfy his thirst). The agent's mistaken belief may explain why he would drink the stuff in the bottle, but we would not say that an agent *has a reason* to drink the gasoline. So, an agent's desire does not give a reason for action if that desire is based upon a false belief, nor will an agent's false beliefs about the means to satisfy a desire give a reason for action (Williams 103).

Williams' third refinement follows from the second. As we see, it is possible for an agent to falsely believe that he has an internal reason to ϕ when he actually has no reason to do so. It is also possible for an agent to be unaware of some reason which actually holds for him (103).

Williams gives a fourth refinement: an agent who engages in deliberation can discover internal reasons which hold for him. Although Williams does not define deliberation, it seems to include a number of rational processes. He states:

A clear example of practical reasoning is that leading to the conclusion that one has reason to ϕ because ϕ -ing would be the most convenient, economical, pleasant, etc. way of satisfying some element in S, and this of course is controlled by other elements in S, if not necessarily in a very clear or determinate way. But there are much wider possibilities for deliberation, such as: thinking how the satisfaction of elements in S can be combined, e.g. by time-ordering; where there is some irresolvable conflict among the elements of S, considering which one attaches most weight to . . . or, again, finding constitutive solutions, such as deciding what would make for an entertaining evening, granted that one wants entertainment. (104)

This array of rational processes not only allows an agent to discover what reasons hold for him, but can actually alter the contents of an agent's subjective motivational set. Desires can be lost through deliberation. For example,

upon discovering that a desire was based upon a false belief, an agent might no longer feel the pull of that desire. Similarly, new desires can be formed. For example, upon reflection, an agent may realize he prefers chess to checkers, given that the former allows for greater variety in play. But in all cases of deliberation, we deliberate from our existing subjective motivational set—the formation of any desire must be related to previous desires.

This fact leads to a relativistic account of internal reasons. We can imagine cases where an agent's subjective motivational set is so constituted as to preclude using deliberation to arrive at some particular desire (e.g., the desire to help others, the desire to join the military). If an agent does not have the relevant desire for ϕ -ing, and cannot get the desire through deliberation, then the agent cannot have a reason to ϕ . For example, if an agent lacks a desire to donate to charitable organizations, and cannot deliberate to such a desire, then he simply has no reason—moral or otherwise—to donate to such organizations. So, the reasons which hold for one person in a particular instance would not necessarily hold for another person in the same circumstances.

Smith on Being Fully Rational

According to Smith, philosophers such as Bernard Williams suggest that there is some connection between what an agent has a reason to do in a particular circumstance and what he would do if he were fully rational (109). Smith's account of being fully rational largely resembles Williams' view of internal reasons. Indeed, he agrees with Williams that for an agent to be fully rational the following conditions must be met (112):

- (i) The agent must have no false beliefs.
- (ii) The agent must have all relevant true beliefs.¹
- (iii) The agent must deliberate correctly.

That said, Smith also notes some weaknesses of Williams' theory. For example, Williams' view cannot necessarily deal with the effects of emotions like anger—"unless some such constraint is supposed to be presupposed by condition (iii), the condition of correct deliberation" (113). But Smith is more concerned with what he considers a greater failing of Williams' view. Williams' account of deliberation leaves out a very important role of our

¹ Although Smith insists on this, it is not entirely clear that Williams requires anything more than that the agent have no relevant false beliefs. Of course, that position may or may not be equivalent to having all relevant true beliefs. But either way, we can at least treat (ii) as an additional refinement proposed by Smith, if not as one that would be welcomed by Williams.

deliberative processes: to find out if our desires are systematically justifiable and as far as possible to make them so (Smith 114). Smith writes:

What do I mean when I say that we sometimes deliberate by trying to find out whether our desires, as a whole, are systematically justifiable? I mean just that we can try to decide whether or not some particular underived desire that we have or might have is a desire to do something that is itself non-derivatively desirable, and that we do this in a certain characteristic way: namely, by trying to integrate the object of that desire into a more *coherent* and *unified* desiderative profile and evaluative outlook. (114)

We can imagine a large number of independent, underived elements in our subjective motivational set. Some of these elements might have general application (e.g., a love of cats), while some might have a very specific application (e.g., a fondness for my sister's bearded dragon). This act of deliberation can be seen as giving structure to these elements by reducing particular desires to more general ones. Similarly, when we find *ad hoc* desires that simply cannot be systematized we will be inclined to discard them (Smith 115). All of these deliberative acts are directed toward a more coherent and unified set of desires.

Smith believes that ideally these sorts of deliberative, systematizing acts will result in some convergence in fundamental, general desires held by people. Or, more accurately, Smith argues that part of being a rational creature engaging in systematic justification is attempting to systematize desires in such a way that they will converge. He writes:

All possible rational creatures would desire alike as regards what is to be done in the various circumstances they might face because this is, *inter alia*, what defines them to be "rational." Part of the task of coming up with a maximally coherent and unified set of desires is coming up with a set that would be converged upon by other rational creatures who too are trying to come up with a maximally coherent and unified set of desires. (118)

This account would give non-relativistic internal reasons, and Smith holds that such an account is more plausible than any relativistic account (such as the one offered by Bernard Williams). It is not my concern to oppose Smith's later arguments regarding Williams' account, so I will not detail them here. Rather, I wish to argue that systematic justification is not a necessary feature of rational deliberation.

Varieties of Systematic Justification: Non-contradiction

The term “systematic justification,” as ordinarily used, seems to have two senses. The first sense deals with avoiding conflict among our various desires and is suggested by analogy with systems of theoretical reasoning—we avoid contradictions in our reasoning, and we should similarly avoid them in our desires.² The second sense deals with placing, where possible, specific desires under more general desires, thereby giving greater unity and coherence to our system of desires. This sense too may be suggested by analogy to theoretical reasoning—a scientific theory which gives a unified account for a number of phenomena is generally considered more satisfactory than a collection of unrelated explanations of events.

The first sense of systematic justification is accommodated by Williams’ view of deliberation. We might imagine a conflict of desires; for example, I want to do my homework tonight, but I also want to play Scrabble. As Williams notes, I have various deliberative options to resolve the conflicts—time-ordering, assigning relative weights to desires, etc.—but no matter which way I actually do it, the conflict is resolved. In this particular case, I would probably assign greater weight to my desire to do homework, and forgo Scrabble for the evening.

One might object that both desires are still present in my subjective motivational set, and they are still in conflict—I have simply chosen one action over the other. This may be the case, for example, if I still want to play Scrabble despite having chosen to do my homework. But I would reply that although both desires are present and pull me different ways (to speak figuratively), they are not *contradictory*—the analogy to theoretical reasoning simply does not hold here. But why would the analogy seem persuasive? Well, contradictory beliefs cannot both be true, and contradictory desires cannot both be satisfied—this resemblance might suggest other similarities. Contradictory beliefs are damaging because they prevent correct thinking (and at least one of the beliefs is false); but contradictory desires? They do not prevent correct desiring nor acting: I do my homework despite the conflicting desire to play Scrabble. Perhaps, if Williams’ theory did not provide a means for me to choose one desire over the other, then we could call the desires contradictory, as their conflict would prevent my choice and action. But as things are, my desire to play Scrabble is at worst a nuisance, given that I will not satisfy the desire this evening.

Still, one might insist that as long as two desires cannot both be satisfied, they are contradictory, and that it would be preferable to have only

²To be fair, Smith does not really deal with this first sense of “systematic justification.” I only treat it in order to show that when Smith speaks about “unity” and “coherence” in a system of desires, he cannot be using these terms to mean “free from contradictions.”

one of them. In this example, it *may* be preferable to have only one desire—it would probably be easier to do homework if I were not tempted to play a board game. But this is not always the case. For example, suppose an agent has five dollars. A representative of the Red Cross approaches him at the same time as someone from the March of Dimes. The agent desires to donate five dollars to each organization, but he cannot, for he does not have ten dollars. Engaging in deliberation, the agent might donate to one organization or the other (by assigning greater weight to one of them); he might even donate some portion of the five dollars to each, etc. But suppose that he decided to give all the money to the Red Cross. Would it be *preferable* for him to not desire to give five dollars to the March of Dimes? It seems that it is not: if we have any preference, we would say that it is better that he desire to give fully to both, even if he cannot do so. The same is true for any other decision he would have made. And the same might be true for my example above—maybe it is not a bad thing that I still desire to play Scrabble. For example, it might reflect laudable interest in socializing with friends.

This highlights an additional problem with calling this sort of conflict a “contradiction.” Conflicts of desire arise easily, even in a system such as Smith’s. Take the desire to donate to charity—I presume that this would be one of the desires that Smith’s variety of a perfectly rational being would have. As seen above, this desire easily leads to a conflict, given that we can only donate a finite amount. Similar conflicts arise from other general desires, especially when we consider them together: the desire for leisure, the desire to raise a family, the desire to help others, the desire to succeed in one’s career, etc. Often the satisfaction of one will leave the others—to some degree—unsatisfied. But we should not be anxious to be rid of the others in these cases. These conflicts of desire arise far too often in normal practical reasoning for us to consider them analogous to contradictions in theoretical reasoning.

Varieties of Systematic Justification: Generalization

Let us consider the second, stronger sense of systematic justification, which deals with placing specific desires under more general desires. I find it difficult to imagine significant cases where I would just *happen* to possess all the elements of a more general set, thus allowing me to generalize. I also suspect that in most cases where the particular elements required for generalization are all present, this is only because the general element was already present. Let me illustrate these claims with examples dealing with the elements of preference in the subjective motivational set.

For example, I might find both Mrs. Jones and Mr. Jones agreeable (so I have reasons to spend time with them), and therefore be able to

generalize to finding the Joneses agreeable. But this sort of case has a very limited scope: it can only occur when I am familiar with all members of the set in question. Moreover, it does not seem to really add anything new to my subjective motivational set; nor does it add anything to my reasons for action which was not already present when I regarded the persons individually. If this is what is meant by generalization, then Williams' view of deliberation should be able to produce the same results.

The majority of cases, however, would probably look something more like what follows. I happen to like greyhounds, poodles, and Chihuahuas, and a great number of similar things. This suggests that I can generalize to say that I like dogs. But can I? Must I? Strictly on the basis of rationality? If there is some antecedent preference, then it might be the case that I can make this generalization as I become aware of the antecedent preference ("You like furry things, don't you? And things that lick you, right?" "Yes." "Well, then you'll like all dogs, because they're all furry and will lick you!"). But again, this is not a case of subsuming underived desires under a more general desire. Also, this sort of generalization could easily be accomplished by Williams' model of deliberation.

What we need, if we are to get results different from those Williams provides, is a case where there are independent, underived preferences, but which preferences can be generalized in such a way as to create a new desire with results different than those which would follow from the antecedent desires alone. In the dog example, we need a case where I just *happen* to like greyhounds, poodles, and Chihuahuas—and not for some general reason, that they are cuddly or soft or faithful or anything else from which I should derive my preference. I like them—I simply like them—and "that's the way I am." Or, perhaps a slightly weaker condition is enough for Smith's view—I like them, but I like each for a different reason: greyhounds because they are strong, poodles because they are elegant, and Chihuahuas because they can fit in a purse.

In this sort of scenario, I may very well deliberate to the more generalized "I love dogs," and this general desire does give results different from the antecedent desires taken on their own—it is a more inclusive preference, dealing also with species which I have not yet met (species which may not be strong, elegant, nor purse-sized). And had I not made the generalizing step, there might have been some member or group of members of the set with whom I have not yet made acquaintance, but which I would have disliked—say, for example, Scottish terriers. So, generalizing in this manner can give a different result than if rationality did not require that the agent systematize his desires where possible.

I said that I "may very well" deliberate to the generalized desire, but this is not strong enough for Smith. If his argument is to hold, it must

be shown that all fully rational beings *would* deliberate to that desire—in other words, there should be something about rationality which brings persons to generalization in these sorts of circumstances. So I ask, *must* a fully rational person deliberate from those antecedent preferences to the more generalized form?

While it is probably true that many rational persons *would* make the generalization, this fact has nothing to do with their rationality, *per se*. The problem is similar to that of generalizing universal truth based upon empirical experience of particulars—many cases of induction do not seem to be required by logic or rationality alone. Rational persons need not make the generalization if they are not so inclined. Indeed, many such generalizations are later seen to be unfounded. For example, although empirical experience long suggested that all swans were white, this generalization was later found to be false—and it would be odd, to say the least, if we insisted that rationality *required* Europeans to make the false generalization that “all swans are white,” only to find that this was false. Similarly, although one might have affection for every species of dog that one has met, rationality does not require that we generalize this affection for all dogs.

Why then, do many—but not all—rational agents choose to generalize their affection in some of these situations? It seems to be because of their beliefs regarding generalization and when to apply it. But they cannot have arrived at absolute certitude regarding these beliefs—or in other words, these beliefs are not dictated by rationality alone. An individual might generalize, if so inclined, for that is not contrary to reason; but the individual might also choose not to generalize. It is not irrational to use induction in some cases but not others. So, a person can be fully rational and yet refrain from generalizing at all or in particular instances.

This seems right. Smith writes that “exhibiting coherence and unity is partially constitutive of having a systematically justified, and so rationally preferable, set of desires, just as exhibiting coherence and unity is partially constitutive of having a systematically justified, and so rationally preferable, set of beliefs” (115). But why should we value coherence and unity? If by these terms we only mean “non-contradiction,” then his statement is correct, but not terribly illuminating—for, unlike beliefs, desires cannot be contradictory. Similarly, if we only value coherence and unity instrumentally, insofar as they promote non-contradiction in both theoretical and practical reasoning, then, once again, that statement is correct. But if this is all that Smith means by “coherence” and “unity,” then let us note that Williams’ deliberation can resolve cases of debilitating “contradiction.”

To make his argument strong enough for a non-relativistic account of internal reasons, Smith must take “coherence” and “unity” to mean

something more than simply non-contradiction. They must represent some value which will require that we subsume specifics under generals where we can. This value will prefer—whenever possible—to give a single, underlying desire behind our many reasons for action rather than multiple desires. There are analogous values in theoretical reasoning, such as one which prefers a monistic cosmology to a pluralistic one, or prefers reducing the electromagnetic and weak nuclear forces to a single electro-weak force. But in these other fields, it is by no means clear that a rational person *must* assent to valuing that reduction simply because it is a reduction. Indeed, especially in the example of cosmology, a rational person could just as well value a view which does *not* reduce existence to one thing (or type of thing). The point is that not all rational persons need be theoretical systematizers. The same holds for practical reasoning: not all rational persons need be practical systematizers. Although Smith himself may value coherence and unity in this stronger sense (and we might as well, if we so desire), we can still be rational without valuing them in that way. An agent can rationally have *ad hoc* beliefs, provided they do not contradict the agent's other beliefs. Similarly, an agent can rationally have *ad hoc* desires, provided the agent has some means for resolving conflicts with other desires.

In conclusion, Williams' account of deliberation can provide coherence and unity in the weaker sense—that is, it can avoid self-contradiction and motivational incapacity. And although it does not require the stronger kind of coherence and unity, a rational person need not value that sort of coherence and unity. Therefore, there can be fully rational persons who nonetheless fail to deliberate in the systematizing manner that Smith requires for his arguments about the non-relativity of reasons.

Works Cited

- Smith, Michael. "Internal Reasons." *Philosophy and Phenomenological Research* 55 (1995): 109–31.
- Williams, Bernard. "Internal and External Reasons." *Moral Luck: Philosophical Papers, 1973–1980*. Cambridge: Cambridge UP, 1981.

Works Consulted

- Clark, Philip. "Velleman's Autonomism." *Ethics* 111 (2001): 580-93.
- Cordner, Christopher. "Ethical Necessity and Internal Reasons." *Philosophy* 76 (2001): 541-60.
- Hollis, Martin. "Jim and the Indians." *Analysis* 43 (1983): 36-39.
- Hooker, Brad. "Williams' Argument Against External Reasons." *Analysis* 47 (1987): 42-44.
- Korsgaard, Christine M. "Skepticism About Practical Reason." *The Journal of Philosophy* 83 (1986): 5-25.
- McDowell, John. "Might There Be External Reasons?" *Mind, Value and Reality*. Cambridge, MA: Harvard UP, 1998.
- Millgram, Elijah. "Williams' Argument Against External Reasons." *Noûs* 30 (1996): 197-220.
- Quinn, Warren. "Putting Rationality in its Place." *Virtues and Reasons: Philippa Foot and Moral Theory: Essays in Honour of Philippa Foot*. Oxford: Clarendon Press, 1995.
- Taylor, Craig. "Moral Incapacity." *Philosophy* 70 (1995): 273-85.
- Velleman, J. David. "The Possibility of Practical Reason." *Ethics* 106 (1996): 694-726.
- Wiland, Eric. "Good Advice and Rational Action." *Philosophy and Phenomenological Research* 60 (2000): 561-69.

