Conative Transcendental Arguments

and the Question Whether There Can Be External Reasons

**Abstract**

A characterization of transcendental arguments is proffered, whereby they yield conclusions about how things are *via* intermediate conclusions about how we must think that they are. A variant kind of argument is then introduced. Arguments of this variant kind are dubbed ‘conative’ transcendental arguments: these yield conclusions about how it is desirable for things to be *via* intermediate conclusions about how we must desire that they are. The prospects for conative transcendental arguments are considered. It is argued that, although they can never be of practical use, they may nevertheless be of use in dissolving certain applications of the debate—initiated by Bernard Williams—about whether anyone can have an ‘external’ reason to do anything, that is to say a reason that is not grounded in some desire of the person’s, in a suitably broad sense of ‘desire’. The relevance of conative transcendental arguments to this debate is that they highlight desires that we cannot help having and with respect to which the debate lacks any suitable focus. In the final section of the essay five conative transcendental arguments deriving from the work of five moral philosophers are briefly considered.

**1. Transcendental arguments**

One fairly standard way of characterizing transcendental arguments would be this. Transcendental arguments are arguments of the form:

(1) *p*

(2) It would not be possible that *p* if we did not think that *q*

 (3) We must think that *q*

 (4) It is true that *q*,

where what replaces '*p*' is something about how we represent reality, either something minimal—such as that we have self-conscious experiences, or that we exercise concepts, or that we make assertions—or something more restrictive, such as that we do these things concerning some specified subject matter. By no means everything that has been given the label 'transcendental argument’ fills this bill: the term has no single accepted usage. But the characterization is standard enough to capture a significant common core in the various uses of the term, and I shall appropriate it as a (stipulative) definition of my own use.[[1]](#footnote-1)

On this construal, there are certain objections to which any transcendental argument is vulnerable. There is the objection that unless (1) is interpreted in such a way that it carries question-begging presuppositions about what exactly we achieve in representing reality as we do, there is no reason to accept (2). There is the objection that (2) can in any case never be known, because there is never any ruling out our making (1) true in some hitherto unimagined way that does not involve our thinking, or even being able to think, that *q*. There is the objection that the modality involved in (2) and (3) is unclear, and that the argument as a whole cannot survive any clarification of it; more specifically, perhaps, that the modality involved in (2) and (3) is a conflation of conceptual and psychological modalities, and that there is no satisfactory way of disentangling these. There is the objection that the step from (2) to (3), or perhaps the double step from (2) to (4), confuses the necessity of a hypothetical with the necessity of its consequent. There is the objection that—Moore's paradox notwithstanding[[2]](#footnote-2)—the final step from (3) to (4) is invalid; or, a little less harshly, that the final step from (3) to (4) needs the support of some highly questionable metaphysical assumption, such as verificationism or idealism.[[3]](#footnote-3)

Different transcendental arguments are obviously differently threatened by these objections, and generate different sorts of counter-objections. I shall not now try to answer the question whether any transcendental argument survives all the objections. My purpose is rather to explore an unfamiliar variation on this familiar theme. I want to consider the prospects for what I shall call 'conative' transcendental arguments.

**2. Conative transcendental arguments**

By 'conative' transcendental arguments I mean arguments whose form differs from that of (ordinary) transcendental arguments in three respects.

The first is that 'think' in (2) and (3) is replaced by 'desire'. (This results in a pair of schemata that have, for the most part, hopelessly unidiomatic instances. However, for an argument to be of this form it is not necessary that its actual wording conform to this pattern, only that the wording of some regimented paraphrase do so.) 'Desire' here is to be understood as a quasi-technical term denoting any conative state. Examples of desiring that *q* are hoping that *q*, wanting it to be the case that *q*, wishing it were the case that *q*, having a whimsical urge to make it the case that *q*, and thinking it only right and proper that *q*.[[4]](#footnote-4)

The second respect in which the form of conative transcendental arguments differs from that of their non-conative counterparts is that 'true' in (4) is replaced by 'desirable'. ‘Desirable' here is to be interpreted in a similarly plastic way to 'desire', the idea being that the replacement for (4) should stand to the replacement for (3) in an analogous relation to that in which (4) stands to (3): a relation, very roughly, of vindication, Just as its being true that *q* means that we are right to think that *q*, so too its being desirable that *q* means that we are right to desire that *q*.[[5]](#footnote-5)

The third and final respect in which the two forms of argument differ is that the original restriction on '*p*', namely that what replaces it is something about how we represent reality, is dropped. It would be natural to insist on a new restriction, namely that what replaces '*p*' is something about how we *desire* reality to be. Conative transcendental arguments would then involve conation, so to speak, all the way down. I shall not insist on any such restriction however. My interest in conative transcendental arguments spans the different kinds of input they can have. I shall deliberately leave open what kind of claim the first premiss is. In particular, I shall allow '*p*' to be replaced by something that is neither (just) about how we represent reality nor (just) about how we desire reality to be but (however minimally and however indirectly) about both: for instance, 'We are rational agents.' It is worth noting, though, that given the various restrictions that might be imposed on '*p*', there is scope for a finer-grained classification of arguments than I am now effecting.[[6]](#footnote-6)

These three differences issue in the following definition of a conative transcendental argument. A conative transcendental argument is an argument of the form:

(1c) *p*

(2c) It would not be possible that *p* if we did not desire that *q*

 (3c) We must desire that *q*

 (4c) It is desirable that *q*.

**3. Good conative transcendental arguments**

It is clear that, just as there are certain objections to which any transcendental argument is vulnerable, so too there are certain objections to which any conative transcendental argument is vulnerable. Indeed the objections listed above all have immediate analogues. Most obvious, and most significant, is the objection that the final step from (3c) to (4c) is invalid.

This objection calls to mind Moore's celebrated reproof of the step in Mill's argument concerning which he said, 'The fallacy in this step is so obvious, that it is quite wonderful how Mill failed to see it.’[[7]](#footnote-7) What Mill had said was:

The only proof capable of being given that a thing is visible, is that people actually see it. The only proof that a sound is audible, is that people hear it: and so of the other sources of our experience. In like manner, I apprehend, the sole evidence it is possible to produce that anything is desirable, is that people do actually desire it.

Mill's step is a fairly close cousin of the step from (3c) to (4c), And, *prima facie* at least, both steps do appear to involve the kind of mistake to which Moore was voicing his well-known opposition, the kind of mistake to which he famously gave the name 'the naturalistic fallacy'.[[8]](#footnote-8) Much has been written about just what kind of mistake this is, and about whether Mill was really guilty of it. He had after all prefaced these comments with a reminder that 'questions of ultimate ends do not admit of proof',[[9]](#footnote-9) and it is significant that in the comments themselves he had spoken not of the sole 'proof' it is possible to produce that anything is desirable, but of the sole 'evidence'.[[10]](#footnote-10) This relates to points that I shall be raising later, when trying to view conative transcendental arguments in a more favourable light. But whatever verdict we give on Mill's step, the fact remains that the step from (3c) to (4c) does look likely to be the weak spot in any serious conative transcendental argument.[[11]](#footnote-11)

Before I proceed, and as a preliminary to trying to make this final step look more reasonable, I shall address another of the objections to which any conative transcendental argument is vulnerable (an objection that is likewise an analogue of one of those listed earlier, to non-conative transcendental arguments): the objection, namely, that the necessity of a hypothetical is being confused with the necessity of its consequent. The objection is that the 'must' in (3c) registers the conditional necessity of our desiring that *q*—conditional on its being the case that *p*—but is then treated as if it registered an unconditional necessity; or rather, that *unless* it is treated as if it registered an unconditional necessity, the final step from (3c) to (4c) is beyond redemption. For, the objection runs, however arguable the inference from our being unable to help desiring something to its being desirable, there is nothing to be said in support of the inference from our merely happening to desire something to its being desirable.

How might this objection be met?

Well, clearly, if it were *necessary* that *p*, the objection would lose its force: the necessity of the hypothetical would then yield the necessity of the consequent. The objection is likely to be met, in any given case, by arguing that the first premiss is a necessary truth.

However, the necessity here—and by 'here' I mean in the original objection as well as in this parry—must itself be conditional. There cannot be any *absolute* necessity about our desiring that *q*. There is no absolute necessity about our existing at all. And the objection was not that nothing less than *this* would prevent the final step from (3c) to (4c) from being beyond redemption.[[12]](#footnote-12)

What then will the necessity be conditional on?

That will vary from argument to argument, Of greatest rhetorical interest will be cases in which the necessity is conditional on something that is in turn a necessary condition of our engaging with the argument in the first place, such as our being able to make judgements about what is desirable, our being rational, or the like. The most direct way for the first premiss to instantiate this necessity is by being, quite simply, a claim to the effect that this condition holds: that we are able to make judgements about what is desirable, that we are rational, or whatever it may be.

Suppose then that it can be shown that the first premiss does indeed instantiate such necessity, and that the necessity is transmitted to our desiring that *q*. Then certainly the final step from (3c) to (4c) will look more reasonable. (This is indirectly related to the point that, although the desirability of a thing is not the same as the thing's capacity to be desired, the former does require the latter.)[[13]](#footnote-13) The final step will look more reasonable, I say; but still not indisputable. What further considerations can be brought to bear on it?

At this point it helps to consider two comparisons between conative transcendental arguments and their non-conative counterparts.[[14]](#footnote-14) One of these comparisons is to the advantage of conative transcendental arguments. One is to their disadvantage.

The comparison that is to their advantage concerns a basic instinctive realism that inclines us to doubt the step from (3) to (4) in non-conative transcendental arguments, a realism that inclines us to say: *whatever* the inescapability of our thinking that *q*, still what we think may be false. (Maybe we are constitutionally incapable of acknowledging the truth about this matter.) This worry can perhaps be met by a dose of verificationism or idealism.[[15]](#footnote-15) But then the verificationism or idealism is liable to give rise to a similar worry of its own. That is, the original impulse to realism is liable to remain unchecked. In the conative case, by contrast, there is no such impulse. There is not the same urge to say: *whatever* the inescapability of our desiring that *q*, still what we desire may be undesirable.[[16]](#footnote-16) Our thoughts, we are inclined to say, unlike our desires, answer to what is there *anyway*.[[17]](#footnote-17)

The other comparison, the one that is to the disadvantage of conative transcendental arguments, is as follows. In their case, the conditional element in the necessity is still enough to be a real obstacle to the final inference. For even if what the necessity is conditional on is our very capacity to engage with such arguments, there is room for the question whether it is desirable that we have this capacity—in a way in which there is not room for the question whether it is *true* that we have it. To be sure, a very hardened sceptic may say, while engaging with one of these arguments, that he doubts his own capacity to do so. But it requires a much less hardened sceptic to say, while engaging with one of these arguments, that he doubts the *value* of his capacity to do so. He may suspect that we would all be better off if we could not so much as think about these things, in which case there would obviously be a question about the desirability of that which, given that we can think about them, we cannot help desiring.

With these two comparisons in mind, I venture to draw the following conclusion from the discussion in this section: the first and most basic worry about any conative transcendental argument, the worry that the final step from (3c) to (4c) is invalid, can satisfactorily be met by showing that the first premiss enjoys a conditional necessity, *provided*—and this is the lesson of the second comparison—that what the necessity is conditional on can itself be shown to be desirable. This may look like a convoluted way of saying that whatever is a necessary condition of something desirable is itself desirable. But it is not. And it is as well that it is not. That claim—the claim that whatever is a necessary condition of something desirable is itself desirable—is false, or at best dubious. Conative transcendental arguments take us beyond that claim because of their focus on what we cannot help desiring.

Nevertheless, the force of the proviso is that no conative transcendental argument, as it stands, is sound. At best it is enthymematic, requiring the addition of a premiss guaranteeing the desirability of its input. But if we say that a conative transcendental argument is 'good' when it can be converted into an argument that is sound by the addition of just such a premiss, then at this stage there is no reason to think that there cannot be any good conative transcendental arguments.

**4. Why good conative transcendental arguments can never be of practical use**

Suppose there can. How might they be of use to us?

Well, one note of caution needs to be sounded straight away. Suppose there were some impediment (of an empirical kind, perhaps) to its being the case both that *q* and that *r*. And suppose there were a good conative transcendental argument establishing[[18]](#footnote-18) that it was desirable that *q*. Even so, it might also be desirable that *r*. Indeed, there might be a good conative transcendental argument establishing that it was desirable that *r*. (Who knows but that we cannot help having conflicting desires?)

In a way this note of caution has nothing specifically to do with conative transcendental arguments. Whenever we decide that something is desirable, we must reckon with the possibility that something else, something incompatible, is also desirable, and perhaps more desirable. But in a way the note of caution *is* specially pertinent in the case of conative transcendental arguments. To see why, we need to address some important questions, hitherto shelved, about the interpretation of the schema 'We desire that *q*'. How exactly is this schema to be understood?

Let us leave aside the issue of who 'we' are. (This is non-trivial; but it is not the main point of concern here.) The question we need to address first is whether the desire is a corporate desire, or a desire that each of us has severally. The latter is overwhelmingly the more natural interpretation. In fact it is not obvious what a corporate desire would be.[[19]](#footnote-19) True, there is a natural worry concerning the non-corporate interpretation: an awful lot seems to be required for each of us to desire that *q*. For instance, whatever attractions a thing may have, there is a certain basic level of rationality, attentiveness, knowledgeability, and so forth that any given individual must attain in order to be guaranteed an active interest in pursuing it. However, we should not forget the broad notion of 'desire' involved here—much broader than the notion of having an active interest in pursuing something—nor the narrow notion of who 'we' are that may also be involved. Once these two things have been taken into account, the claim that each of us desires that *q* will seem less demanding. Let us take for granted, then, that the schema does require each of us to desire that *q*.

Now prima facie '*q*' must be replaced by something that is true or false independently of this particular linguistic context. That is, it must not involve an anaphoric cross-reference. Thus consider the following instance of the schema:

We desire that we are happy.

This must, it seems, be understood in the sense that each of us desires that each of us is happy, not in the sense that each of us desires that *he or she* is happy. The reason for this is that '*q*' also appears in the schema 'It is desirable that *q*'. A cross-reference in the original case would lead, it seems, to incoherence in this. Thus while it makes sense to say that each of us desires that he or she is happy, it makes no sense (or no self-standing sense) to say that it is desirable that 'he or she' is happy.[[20]](#footnote-20)

There is now a problem however. For if there are any desires that we cannot help having, then at least some of them, and arguably all of them, will be desires of the very kind just precluded, desires that each of us has concerning himself or herself ('desires *de se*', as we could call them). If good conative transcendental arguments are to exploit these desires, perhaps indeed if there are to be any good conative transcendental arguments, then we had better reassess the considerations that have brought us to this point.

One possibility would be to revoke our assumption that the schema 'We desire that *q*' requires each of us to desire that *q*. We could specially construct a notion of corporate desire, whereby a group has the corporate desire that *q* if and only if at least one of its members desires that *q*; and we could understand the schema in accord with this. Thus if each of us desires that he or she is happy, then what *we* desire, in this specially constructed corporate sense, is that I am happy, that you are happy, that he is happy, and so forth, one such proposition for each of us.[[21]](#footnote-21)

The problem with this is that such gerrymandering would serve only to put further pressure on the final inference from (3c) to (4c): the mere fact that *one* of us cannot help desiring that *q* is a much weaker reason for concluding that it is desirable that *q*.[[22]](#footnote-22) This pressure would be relieved if the desirability in question could be construed as desirability relative to the individual concerned. But in that case there would be a far more satisfactory alternative. In fact there is a far more satisfactory alternative anyway It is this: to construe (4c) as an abbreviation for:

(4c\*) It is desirable for us that *q*,

allowing '*q*' to contain an anaphoric cross-reference after all. 'For us' here acts as a universal quantifier whose scope is the entire schema; but it also picks up on any relativization to individuals that may be involved. Thus suppose it makes sense to say that it is desirable for an individual that he or she is happy. Then one instance of the schema (4c\*)—or of (4c), if understood as an abbreviation for (4c\*)—would be:

It is desirable for us that we are happy,

meaning: for each of us, it is desirable for him or her that he or she is happy. But of course, this raises the spectre that what can be shown to be desirable in my case is in conflict with what can be shown to be desirable in your case, and that each of these is in conflict with what can be shown to be desirable in his case, *et cetera*: imagine, for instance, that what each of us cannot help desiring is that he or she *is happier than everyone else*. Here at last we come back to the point of concern from which this discussion took off. For while it would be coherent to claim that it is desirable for us that we are happier than everyone else (where the quantifier, remember, has widest scope—so this would mean that, for each of us, it is desirable for him or her that he or she is happier than everyone else, or perhaps even desirable *tout court* that he or she is happier than everyone else), this claim would nevertheless clearly involve the desirability of incompatibles, or the relative desirability. of incompatibles. Not that there is any reason to suppose that each of us does desire that he or she is happier than everyone else, still less that each of us cannot help desiring this. But once conative transcendental arguments are understood on this model, as it seems they must be, then it is a real possibility that the things which are shown to be desirable, or relatively desirable [[23]](#footnote-23) are things which are (at least empirically) not compatible. [[24]](#footnote-24)

The seemingly bleak prospect to which the discussion so far points, then, is this. No good conative transcendental argument can ever be of practical use; or, a little more precisely, no good conative transcendental argument can ever constitute a piece of practical reasoning. Plainly no such argument can ever instil in us the desire for that which it shows to be desirable. This is for the reason (apart from any other) that the desire is one which, if the argument really is good, we must have anyway. But nor, we now see, can such an argument give us a decisive reason to implement our desire. Not only might the thing shown to be desirable conflict with other things that are desirable, it might conflict with other things that can be shown to be desirable by the very same argument.

True, it would be a little premature to conclude at this stage that no good conative transcendental argument can ever be of practical use. There are some important intermediate possibilities, that is to say intermediate possibilities between our desiring something and our seeing that we have a conclusive reason to implement our desire. One of these is the possibility of our coming to *recognize* that we have the desire. Another is the possibility, precisely, of our concluding that the thing is desirable. There can obviously be practical significance in our doing either of those things. Even so, I suspect that no good conative transcendental argument *can* ever be of practical use. For instance, concerning the second of these possibilities, I suspect that unless we already acknowledge the desirability of that which the argument shows to be desirable, then nothing in the argument itself has the jurisdiction to make us put more faith in the goodness of the argument than in the falsity of its conclusion. In particular, I suspect that there is never any foreclosing doubts about the suppressed premiss, the premiss guaranteeing the desirability of the argument's input.

However, I do not think that this need be viewed as a 'bleak prospect'. Suppose we do already acknowledge the desirability of that which the argument shows to be desirable. There is still a significant rôle that the argument can play. It can provide a *justification* for our regarding the thing as desirable, a *deduction* to use Kant's word.[[25]](#footnote-25) And here at last we begin to glimpse the most exciting potential that conative transcendental arguments have. They have the potential to dissolve certain applications of the debate about whether there can be external reasons.

**5. Why good conative transcendental arguments may be of use in dissolving certain applications of the debate about whether there can be external reasons**

This is a debate initiated by Bernard Williams. The question is whether there can be reasons for anyone to do anything which are not grounded in some element of the person's 'subjective motivational set', that is in some desire of the person's (in the broad sense of 'desire' that I have appropriated in this essay). External reasons are defined to be reasons of this kind. Williams's own view is that there cannot be such reasons. Call this view internalism, and the opposite view externalism.[[26]](#footnote-26)

Now in order for somebody's reason for doing something to be an external reason, it is not necessary that she actually lack any corresponding desire. (She may want to do what she has a reason to do anyway.) But it is clearly sufficient. And McDowell, in a recent defence of externalism, considers someone in just such a position, that is someone with a reason for doing something but without any corresponding desire. He asks how she could come to acquire such a desire. Not by correct reasoning, he admits. But he does not see why, as an externalist, he should be embarrassed by this. He is quite happy with the alternative, namely that she could only come to acquire the desire as a result of something that 'would not count as [her] being swayed by reasons'; and as other possibilities, he cites, in partial echo of Williams, her 'being persuaded by moving rhetoric', her receiving 'inspiration', and her undergoing 'conversion'. [[27]](#footnote-27) Not that this is an exhaustive or even a fully representative list of all the possibilities there are: it is not intended to be. Presumably her receiving a blow to the head with a hammer, or her taking drugs, *could* give her a desire of the sort in question.[[28]](#footnote-28) The possibilities cited are the ones that would merit a certain kind of approval (approval that would include, but not be exhausted by, approval of their outcome). Or, to put the point with an element of caricature, the phrases 'moving rhetoric', 'inspiration', and 'conversion' are 'hooray'-phrases used to register a certain kind of approval. Other people might register disapproval of the same processes by talking rather of corruption or manipulation or brainwashing. Not that this, in itself, is any threat to McDowell of course. After all, he could say that anyone who disapproves of someone's coming to acquire a desire corresponding to an external reason, through a process that in fact merits approval, is simply in error. However—and this is something that McDowell would himself be the first to concede—he has no Archimedean point from which to justify saying this. He has no Archimedean point from which to arbitrate between those who have the 'right' desires and those who do not. And this immediately suggests the possibility that the same applies to his talk of external reasons; that the phrase 'external reason' is likewise a 'hooray'-phrase, used by externalists to register a certain kind of approval of a certain kind of behaviour; and that there is no Archimedean point from which to arbitrate between those whose desires mean that they approve in this way of one kind of behaviour and those whose desires mean that they approve in this way of some quite different kind of behaviour. Very crudely, it looks as though externalists, when they say that someone has an external reason to do something, mean only that the person's doing the thing would be desirable,[[29]](#footnote-29) there being nothing to privilege their conception of what is desirable over its various rivals. But if this is so, then it surely constitutes a victory for internalism. For internalists can equally say that the person's doing the thing would be desirable.[[30]](#footnote-30) They may also point out that there is already a rich vocabulary of 'hooray'-phrases that can be applied to her doing the thing, as indeed there is a rich vocabulary of ‘boo'-phrases that can be applied to her insofar as she lacks any desire to do it; and that, if talk of external reasons is somehow meant to *enrich* this vocabulary, then we are still owed an account of how. Thus Williams envisages a man of whom it may be said that he is 'ungrateful, inconsiderate, hard, sexist, nasty, selfish, brutal, and many other disadvantageous things'. He then raises the question, 'If [the form of words “He has a reason to be nicer”] is thought to be appropriate, what is supposed to make it appropriate, as opposed to (or in addition to) all those other things that may be said?' [[31]](#footnote-31)

To be sure, we can imagine an externalist replying, 'Nothing. That is, nothing is supposed to make this form of words appropriate *as opposed to* (or *in addition to*) all those other things that may be said. For this man to be ungrateful, inconsiderate, and all the rest *is* for him to have a reason to be nicer.'[[32]](#footnote-32) Such an externalist may also protest against our calling this a victory for internalism. Surely it can just as well be viewed as a victory for externalism? Neutrals, meanwhile, may begin to suspect the debate of being a terminological one.

There is, however, good reason to call this a victory for internalism *if it really is the end of the debate*. The point is this. According to what has been said so far, externalism grants no substance to the claim that somebody has an external reason to do something that is not grounded in *somebody's* desires—'ours', if not hers. Thus whatever licence there may be to make this claim about her (that she has an external reason to do this thing), it has no purchase on her that cannot be offset by the purchase of an equal and opposite claim that she may make, with just as much right (if not more), by saying, 'I have no reason to do this thing'. If externalism is to secure a victory here, then it needs to do more than show how the actual *words* 'She has an external reason to do this thing' can have a legitimate use. It needs to establish an asymmetry between competing demands that can be placed on those words by that use. Thus the fact that McDowell does not have an Archimedean point from which to establish such an asymmetry is, it would now appear, a real problem for him.

McDowell, of course, will deny that it is a problem. He will say that he does not need an Archimedean point; that all he needs is the point, or the ethical outlook as we might call it, from which things are seen 'aright' .[[33]](#footnote-33) From there, he will say, it is possible to establish the ethical primacy of certain desires over others, and, with it, the required asymmetry. If this is so, then the debate has a further twist and may yet culminate in a victory for externalism. What we have here is a familiar dialectical deadlock, where claims for the supremacy of a certain outlook, from within that very outlook, are pitted against claims for the parity of the outlook with its rivals, from within none of them—or supposedly from within none of them. I do not myself believe that there is a single correct way of resolving all such disputes. Different things need to be said in different debates about different kinds of outlook. In this particular case, it seems to me, the idea that there is a 'right' outlook is spurious, and I think that the ultimate victory lies where it originally seemed to lie, with internalism. But I shall not try to argue for that now.

The point I want to make now is this. However the debate between internalism and externalism is resolved, there is one area in which it comes to nothing, namely, where there are no rival outlooks of the kind just described in the first place; or more to the point, where the existence of such rival outlooks is impossible because the existence of the relevant differences of desire is impossible. Thus suppose there is something we *cannot help* desiring. And suppose we have corresponding reasons for doing things. Then with respect to those reasons the debate between internalism and externalism lacks any suitable focus. In particular, it is vacuous to ask whether we would have had those reasons even if we had not had the desire. Internalists can content themselves by saying that each of us has those reasons *because* he or she has the desire. Externalists can content themselves by saying that each of us would have had those reasons *whatever* desires he or she had had (whatever the make-up of his or her 'subjective motivational set' had been)—the point being that, whatever desires any of us had had, they would have included this one.[[34]](#footnote-34) In fact there are many other characteristically externalist things that externalists can content themselves by saying, things which in other contexts would immediately register their externalism but with which, in this context, internalists need not disagree: for instance, that insofar as any of us refuses to *acknowledge* one of the reasons in question, this can only be because of a failure to see the matter aright. The reasons in question are reasons that we all have. They are reasons that we cannot help having.[[35]](#footnote-35)

It should now be clear why I said that conative transcendental arguments have the potential to dissolve certain applications of the debate between internalism and externalism. Precisely what a good conative transcendental argument would show is that there is a desire that we cannot help having.

However, there are three important caveats that need to be entered straight away if this prospect is not to sound more exciting than it really is. First, 'certain applications' is the operative phrase. There is nothing here to suggest that the debate as a whole can be dissolved. If externalists want to insist that there are cases in which people do lack the desires corresponding to reasons they have, or even that there are cases in which people *could* lack the desires corresponding to reasons they have, then, at least with respect to those reasons, or putative reasons, the quarrel with internalists remains as real as ever. Quite how much impact conative transcendental arguments can have on this debate therefore depends on quite what externalists do want to insist. (As Williams complains, 'one of the mysterious things about the denial of internalism lies . . in the fact that it leaves it quite obscure when [the] form of words [“He has a reason to do this”] is thought to be appropriate. ')[[36]](#footnote-36)

The second caveat concerns the modality involved in any good conative transcendental argument, the sense of 'must' in which the argument shows that we must have a certain desire. I have already talked at some length about this. In particular, I pointed out that the modality must be conditional—on our being rational agents, for instance. But even once it has been settled what the modality is conditional on, there remain questions, concerning which I have said virtually nothing, about what *kind* of modality it is, questions, in effect, about what kind of necessity attaches to (2c). Is it conceptual? Is it metaphysical? Might it even be biological? psychological? sociological?

Now, as we shall see, the second caveat does not depend on any particular answer to these questions. Even so, it is worth digressing briefly to say something about how I would answer them. I have been assuming that the necessity is pretty 'hard'. For example, when I talked about the necessity of (1c) being transmitted to the necessity of (3c), I took for granted that the necessity of (2c) was 'hard' enough to carry the transmission; that it did not, for instance, introduce some contingency about our animal nature that was lacking in the first premiss. Again, when I said that we could satisfy ourselves about the validity of the step from (3c) to (4c) provided that we could satisfy ourselves about the truth of the suppressed premiss guaranteeing the desirability of the argument's input, I took for granted that the necessity of (2c) was 'hard' enough not to introduce some extra sully; that it did not, for instance, turn on some dark inability of ours to act except under conditions of self-deception. (Here one thinks of various familiar paradoxes of utilitarian thinking—paradoxes in which our maximizing happiness depends on our desiring something other than the maximization of happiness, something that may be positively undesirable on a utilitarian conception.) But taking these things for granted, which is something I shall continue to do, has only been a convenience. For clearly there is no effective difference between a 'soft' modality and a 'hard' modality that is conditional on whatever makes the 'soft' modality 'soft'. There is no effective difference, for instance, between an argument in which the necessity of (2c) is 'soft' because it incorporates facts about our psychology and an argument in which the necessity of (2c) is 'hard' but the necessity of (1c) is conditional on those same facts.[[37]](#footnote-37)These two features of a modality—how 'hard' it is and what it is conditional on—can be brought together under the single head of how fine its grain is, where the fineness of grain of a modality is a matter of how many possibilities it embraces. If one modality has a finer grain than another, then there are things that are possible in respect of the former but not in respect of the latter (for instance, there are things that are conceptually possible but not biologically possible) though not vice versa. Thus the 'harder' the modality, the finer its grain; and the more it is conditional on, the coarser.[[38]](#footnote-38)

The second caveat is this. The capacity of any conative transcendental argument to dissolve some application of the debate between internalism and externalism depends entirely on whether the modality involved in the argument is as fine-grained as that which is supposed to give the application its focus. Thus even if there is a desire that we 'must' have, in some relatively coarse-grained sense, internalists and externalists can still disagree about whether we 'must', in some finer-grained sense, have the corresponding reasons. There might, for instance, be a good conative transcendental argument showing that we cannot help having a certain desire insofar as we are rational (or granted that we are rational). But if so, this would settle nothing in a debate about whether the corresponding reasons are reasons even for those who are irrational.[[39]](#footnote-39) At the limit there may even be externalists whose very point is that the necessity of our having certain reasons transcends the necessity of our having any desires. Conative transcendental arguments can do nothing to close the rift between *them* and internalists.[[40]](#footnote-40)  In sum then: conative transcendental arguments can dissolve applications of the debate between internalism and externalism only where the modalities involved are properly aligned.[[41]](#footnote-41)

The third caveat is that, if we actually attempt to put a conative transcendental argument to this kind of use, then we are in danger of simply rehearsing parts of the original debate. Suppose, for instance, that the first premiss of the conative transcendental argument is that we have reasons for our desires. Then the second premiss will be that we could not have reasons for our desires unless, in particular, we desired that *q*. But whether *that* is true may just be a question of whether we have an external reason to desire that *q*.

These three caveats notwithstanding, the prospects for conative transcendental arguments, in this particular application, look favourable. That is to say, if there are any good conative transcendental arguments, then their prospects of being put to effective use in this way look favourable. I am not talking about the prospects of there being good conative transcendental arguments in the first place, about which I have been non-committal.

On this score, it is worth pausing to mention what seems to be an anomaly in my essay. As I pointed out earlier, the main obstacle to supposing that there can be any good conative transcendental arguments lies in concern about their final step. I tried to indicate ways in which this concern might be met. Yet that step appears irrelevant to what we have just been discussing. In other words, it appears irrelevant to what I am now suggesting is the most significant application that such arguments can have. Could we not have saved ourselves a lot of bother, therefore, by considering arguments that stopped, so to speak, at (3c)?

No. The charge of irrelevance is ill-founded. Imagine a sound argument that did stop at (3c). Unless the extra step to (4c) were also valid—or rather, unless the step to (4c) would be valid given the addition of the relevant premiss concerning the desirability of the argument's input—then the necessity at work in the argument would not be of a kind fit for the argument to be applied in this way. It is our not being able to help desiring something *in a way that makes it desirable* which allows for the dissolution of the debate about whether our corresponding reasons are external. Remember: what a good conative transcendental argument supplies is a *deduction*, in Kant's sense, of our having some desire. It shows that we are right to have the desire. An argument which did not show this—an argument which showed merely that we must have the desire (perhaps in a relatively coarse-grained sense of 'must' conditional on our having some feature that was itself *un*desirable)—would not be suitable for these purposes. Any attempt to apply it to the debate between internalism and externalism would straightway succumb to the worry expressed in the second caveat, the worry that the disregarded possibilities in which we lack the desire are precisely the possibilities that are pertinent to the debate.

**6. Are there any good conative transcendental arguments?**

That more or less concludes what I have to say in this essay about the prospects for good conative transcendental arguments. It would be perverse, however, to finish without doing *something* to indicate whether there are any. I shall cite five arguments which I think are interesting and serious candidates, each taken, with modifications, from the work of a notable moral philosopher. In each case I shall do little more than state the argument, which will in turn be a matter of stating what replaces '*p*' and '*q*' in the overall schema. I shall not attempt any defence. That would have to be a large and separate undertaking. In any case, I am not convinced that any of these arguments is good. (My non-commitment has been a result of genuine agnosticism.)

(i) The first argument, unsurprisingly, derives from Kant. Here '*p*' and '*q*' are replaced as follows:

*p*: We are finite rational agents

*q*: We are happy.

The 'we' in the second of these is intended to cross-refer (each of us desires that *he or she* is happy). The species of desire involved is hope. The idea is that it is impossible for finite rational beings to act except in the hope that they achieve happiness. In the specific case of those who act as they ought, and thereby make themselves worthy of happiness, the hope is that they achieve the happiness they deserve. Happiness is thus a good. It is not an *unqualified* good; indeed, in cases where those who enjoy it are not worthy of it, it is an apt object of disapprobation. But it is an indispensable part of the 'highest good', an ideal in which virtue and happiness come together in joint reciprocal maximization. [[42]](#footnote-42)

1. The second argument derives from Bentham. This time '*p'* and '*q*' are replaced as follows:

*p*: We make moral judgements

*q*: Happiness is maximized.

Bentham's defence of this argument effectively rests on two claims: first, that in order to make moral judgements, we need to operate with a moral principle having objective and public criteria of application; and second, that the principle of utility, that is the principle that happiness is to be maximized, is the only principle that satisfies this condition. He does not use this argument as a *proof* of the principle (a proof that it is desirable that happiness is maximized). On the contrary, he claims that the principle is an ultimate principle that neither needs nor admits of proof. What he does do is, precisely, to use the argument as a 'deduction' of the principle, or rather of our use of the principle.[[43]](#footnote-43)

1. The third argument is adapted from an argument of Williams. Here '*p*' and '*q*' are replaced as follows:

*p*: We are rational agents

*q*: We are free.

As in the first argument, the second 'we' is intended to cross-refer (each of us desires that *he or she* is free). Williams's own argument is not designed to show any more than what, as rational agents, we must want, namely to be free (and to have an adequate range of wants on the basis of which to exercise our freedom). He does not himself endorse the final step of the conative transcendental argument, the step to the conclusion that it is desirable (for us) that we are free—basically for reasons implicit in objections that I canvassed above. [[44]](#footnote-44)

1. The fourth argument is one that can be extracted from Wiggins, whose own purpose is to show that, with respect to certain fundamental questions concerning the foundations of morality, 'what makes all the difference between Kant and Hume . . . [has] the width of a knife-edge.' Here '*p*' and '*q*' are replaced as follows:

*p*: We are rational

*q*: We enjoy the solidarity of human beings *qua* human.

This time the second 'we' is not intended to cross-refer (each of us desires that *we* enjoy the solidarity of human beings *qua* human). Wiggins starts from the Kantian idea that, insofar as we are rational, we cannot help but aspire to belong to 'the kingdom of ends, the systematic union of rational beings under common self-legislated rational laws'. This Wiggins describes as 'the solidarity of rational beings *qua* rational'. He then asks rhetorically whether 'one who rests morality on a solidarity of this kind is well placed to dismiss a theory that rests morality in *another solidarity*, … the solidarity. . . of human beings *qua* human. '[[45]](#footnote-45)

1. Fifthly, and finally, Korsgaard offers an argument which she herself calls a transcendental argument. In this case '*p*' and '*q*' are replaced as follows:

*p*: We act rationally

*q*: There is such a thing as our humanity

It makes no effective difference in this case whether or not the 'our' cross-refers, since our humanity is shared. The species of desiring that is involved is *valuing*. 'We desire that there is such a thing as our humanity' is meant to stand proxy for 'We value our humanity', the idea being that our humanity is the source of any reasons we have for how we act and we cannot have those reasons without valuing that source. Now it is obviously a contrivance to represent our valuing our humanity as our having a propositional attitude in this way. Indeed some may think that it is a serious mistake to do so. But even if they are right, Kosgaard's argument is at the very least closely related in form to a conative transcendental argument, and raises essentially the same issues. Most notably, as she expressly acknowledges, it raises the issue of whether, and how, our having to value something ensures that it is valuable. [[46]](#footnote-46)

I am sure that there are plenty of other equally compelling examples. Specially worthy of investigation, I think, are examples in which what replaces '*p*' is something minimal about how we represent reality: [[47]](#footnote-47) 'We have beliefs', say, with '*q*' being replaced by 'Our beliefs are true'. Also worthy of investigation are examples in which what replaces '*p*' is the *same* as what replaces '*q*', so that the conclusion of the argument is at the same time (in effect) the suppressed premiss whose truth is required for it to be good. (An argument of this kind would obviously have no suasive force. But suasion, as I have already urged, is not the point.) Finally, I myself have a particular interest in examples in which what replaces '*p*' is something about our craving for infinitude.[[48]](#footnote-48) But these are all matters to be addressed on another occasion.[[49]](#footnote-49)

1. Cf. Cassam (1987), p. 355; Harrison (1982), p. 211; Lear (1984), pp. 219 ff.; Walker (1978), p. 10; and esp. Stroud (1968), passim. Note: despite the obvious Kantian resonances of the term, and despite the affinities of so much of what has been written about transcendental arguments with Kant's work, it is a term that he himself almost never uses. (The sole occurrence of it in his (1933) is at A627/B655—though see also A 589/B617. At A84—95/ B116—129, he explains his use of the related term 'transcendental deduction': see esp. A85/ B117.) [↑](#footnote-ref-1)
2. Moore's paradox has to do with puzzles such as this: although it is possible for me to think something false, there is an incoherence in my saying, of any given proposition, both that I think it and that it is false (see Sorensen (1988), Ch. 1)). We would be involved in a similar incoherence if we claimed that (3) was true and that (4) was false. [↑](#footnote-ref-2)
3. For exploration of these and other objections, see Cassam (1987); Harrison (1982); and Stroud (1968). [↑](#footnote-ref-3)
4. A phrase that could just as well have been used in place of 'desire' here, were it not so cumbersome, is 'have a pro-attitude to its being the case': cf. Davidson (1980a), pp. 3—4. Note: there are important questions, which I shall not address, about just what count as conative states. In particular, do needs? Here I echo Bernard Williams, who, having famously introduced his own broad notion of an agent's 'subjective motivational set', raises the question of whether the agent's needs belong to this set ([B.] Williams (1981b), pp. 102 and 105). He decides not. In similar vein I shall stipulate that conative states, on my understanding, do not include needs. For like Williams, I am interested in motivating states. And although I am more suspicious than he is about the possibility of non-motivating needs, I certainly do not claim to be able to rule the possibility out. [↑](#footnote-ref-4)
5. Strictly, I should say: ‘right *in that respect*'. (We could be right to think that *q* in respect of its being true that *q*, but wrong in some other respect, for example in respect of our limited evidence. Similarly, *mutatis mutandis*, in the case of its being desirable that *q*.) Note: to say that it is desirable that *q*, on this interpretation, is not to rule out the possibility that we have an overriding reason to ensure that it is not the case that *q*. For while we may be right. to desire that *q*, we may also be right to desire that *r*, where it is not possible that both *q* and *r*; and the latter desire may (rightly) prevail. For amplification of this point, and for further important qualifications concerning the interpretation of 'desirable', see below, §4. [↑](#footnote-ref-5)
6. Furthermore, given the distinction between arguments that proceed to the truth of something *via* our having to think it and arguments that proceed to the desirability of something *via* our having to desire it, we can extract a subsidiary fourfold classification that may be of some interest: into those arguments of this *genre* which, intuitively speaking, proceed from representation to representation; those which proceed from conation to conation; those which proceed from representation to conation; and those which proceed from conation to representation. The first of these are what I am calling transcendental arguments. The second and third are included in what I am calling conative transcendental arguments. The fourth I am not discussing at all, though they raise interesting questions in their own right. [↑](#footnote-ref-6)
7. [G. E.] Moore (1959a), p. 67. The argument in question, partially quoted below in the main text and quoted by Moore on p. 66 of his ibid., is from Mill (1962), pp. 288—289. [↑](#footnote-ref-7)
8. [G.E.] Moore (1959a), p. 10. [↑](#footnote-ref-8)
9. Mill (1962), p. 288. [↑](#footnote-ref-9)
10. Cf. Hollis (1985), pp. 30 – 31, and Skorupski (1989), pp. 285 – 286. [↑](#footnote-ref-10)
11. There does not even seem to be an analogous incoherence to that involved in our claiming both that we cannot help thinking something and that it is not true (see above, n. 2). There seems to be nothing wrong with saying both that we cannot help desiring something and that it is not desirable; nothing, that is, unless the step from (3c) to (4c) is valid after all. See further below. [↑](#footnote-ref-11)
12. How can the necessity involved in the original objection be conditional, given that the objection itself made explicit reference to unconditional necessity? Well, here is an analogy. We can distinguish between the merely conditional necessity of Albert's dying of cirrhosis— conditional on his not stopping drinking—and the unconditional necessity of his dying, period. But the necessity in both cases is conditional on the laws of nature. [↑](#footnote-ref-12)
13. Cf. Thomson (1987) passim. Note: if something's desirability means that it ought to be desired, then the fact that the former requires the latter is a simple corollary of the principle '*ought* implies *can*'. [↑](#footnote-ref-13)
14. A good deal of what I have just said applies as much to non-conative transcendental arguments (whose analogous objection elicits an analogous sequence of parries and ripostes) as to conative ones. Thus it is true in their case too that the objection is likely to be met by showing that the first premiss of the argument enjoys a certain necessity; that this necessity must be a conditional necessity; that, if the argument is to have a certain rhetorical force, then what the necessity is conditional on will be something that is in turn a necessary condition of our engaging with the argument in the first place; and that the first premiss will often instantiate the necessity by being a claim to the effect that this condition holds. Cf. Harrison (1982), pp. 214—216. [↑](#footnote-ref-14)
15. There are two ways in which this might work. The first is by closing the gap between the best view that we can form about a matter and the truth of the matter: cf. Stroud (1968). The second is less direct. It proceeds from the assumption that the gap can be closed anyway, say by appeal to what it is for our view about a certain matter to be a view about *that matter*. Idealism is invoked to explain something which, on this assumption, would otherwise be a mystery, namely that, in some cases, there is only one view we can have. The thought is: it is only with respect to what has an ideal existence that there *is* only one view we can have (or in other words, that there is anything we must think). Cf. [B.] Williams (1973d), p. 128. [↑](#footnote-ref-15)
16. Cf. [B.] Williams (1972), p. 49. Here it is interesting to note that even Thomas Nagel, that most robust of realists, is not prepared to extend his realism to ethics: see Nagel (1986), p. 139. Part of what is going on here, of course, is the celebrated distinction between the two directions of fit of representation and conation: see Platts (1979), pp. 256 - 257; and cf. Anscombe (1957), §32. Note: the broad sense of 'desire' that I am using in this essay must not be allowed to slur over this distinction. One example I gave of desiring that *q*, namely thinking it only right and proper that *q*, is not wholly conative. It is also partly representative. But it is an example of desiring that *q* only insofar as it *is* conative. And only insofar as it is conative do the comments above apply to it. (Insofar as it is representative it comes within the ambit of the impulse to realism.) One thing that the breadth of my notion of desire does do, however, is to accentuate the importance of the necessity in (3c). There may be narrower notions of desire whereby the inference from our contingently desiring something to its being desirable looks more reasonable. But these are not specially related to conative transcendental arguments, and they are not my concern. [↑](#footnote-ref-16)
17. Here again there is an allusion to Bernard Williams: see [B.] Williams (1978), p. 64. On the main point being made in this paragraph, cf. Harrison (1976), p. 42. On what is involved in our thoughts' answering to what is there anyway, see my (1997), where I also try to say something about how we *know* that our thoughts answer to what is there anyway.—A question that is related to these issues, and worth pausing to consider, is this. In desiring something, must one already regard it as desirable—*omne appetitum appetitur sub specie boni*—so that someone who professes both to desire something and not to regard it as desirable is involved in a kind of analogue of Moore's paradox? Surely not (see above, n. 11). In regarding something as desirable one is adverting to an additional objectivity, some kind of claim on the thing's being desired from other than one's own current point of view: cf. [B.] Williams (2006i), pp. 58 – 59. What is true is that one cannot desire something without feeling some impulse, however slight, to regard it as desirable. (This is related to, but different from, the point that G.E.M. Anscombe makes in her famous discussion of wants and desirability characterizations: see esp. her (1957), §37. For Kant's discussion of essentially the same issue, see his (1956), pp. 61 ff.— One way to look at the matter would be this: in terms of any commitment beyond one's current perspective, the relation between 'It is desirable that *q*' and 'I desire that *q*' does not correspond to the relation between 'It is true that *q*' and 'I think that *q*', but rather to the relation between 'It is true that *q*' and 'It *seems* to me that *q*', [↑](#footnote-ref-17)
18. Relative to its being merely good rather than sound, that is. This qualification will be implicit whenever I talk in these terms in the remainder of the essay. [↑](#footnote-ref-18)
19. This is not to say that there could not be such a thing. But it is not *obvious* what it would be. [↑](#footnote-ref-19)
20. Cf. Thomas Nagel's distinction between objective and subjective reasons in his (1970), pp. 90 ff.. Cf. also [B.] Williams (1973f), pp. 260 ff. [↑](#footnote-ref-20)
21. Even here there are familiar complications about '*de se*' desires that I am slurring over: John may desire that he is happy but not desire that John is happy, because he does not realize that he is John. Aficionados will notice further examples of my slurring over these complications at various points in what follows. [↑](#footnote-ref-21)
22. Cf. the discussion above in n. 17 on the objectivity of desirability. [↑](#footnote-ref-22)
23. Henceforth I shall tend to leave this rider about relative desirability implicit. [↑](#footnote-ref-23)
24. It is an important question, which I shall not attempt to address here, how far this discussion applies, *mutatis mutandis*, to non-conative transcendental arguments; and, insofar as it does not or cannot, why not. [↑](#footnote-ref-24)
25. See above, n. 1. Kant's own deduction of the categorical imperative can be found in his (1956), Pt I, Bk I, Ch. I, §I. Here it is interesting to note that the categorical imperative was only ever intended to codify what each of us already knows anyway; see e.g. his (1933), A807/B835. (One way to view non-conative transcendental arguments, incidentally, would be as having a similar justificatory role.) [↑](#footnote-ref-25)
26. See [B.] Williams (1981b). See also his (1995a), and (1995c), pp. 186 – 194, which is a response to McDowell (1995). For a further interesting contribution to the debate, see Hollis (1987), esp. Ch. 6. Note: this use of the twin terms 'internalism' and 'externalism' is importantly different from, and indeed in some tension with, that introduced by Falk, in his (1947 - 48), and subsequently adopted by many other writers, e.g. Nagel in his (1970), Ch. II. See further Dancy (1993), App. I. [↑](#footnote-ref-26)
27. McDowell (1995), p. 72; cf. p 74. [↑](#footnote-ref-27)
28. I say 'presumably'. However, I concede that there is important room for (conceptual) doubt about this. [↑](#footnote-ref-28)
29. The qualification 'very crudely' is important here. There is a clear respect in which externalists mean more than this. For instance, suppose that it is desirable that the person do the thing unintentionally. Then there is no question of this constituting her having any kind of reason to do it. Examples can perhaps also be contrived to indicate a respect in which externalists mean *less* than this. Thus externalists might say that Hitler's mother had an external reason to feed him when he was a baby, although, given subsequent events, it was not desirable that she do so. The issues here are complex. For now, I am content just to advert to them. I do not think that they affect the substance of my argument. [↑](#footnote-ref-29)
30. This too, of course, is subject to the qualifications signalled in the previous footnote. [↑](#footnote-ref-30)
31. [B.] Williams (1995a), p. 39. [↑](#footnote-ref-31)
32. Cf. McDowell (1995), pp. 75 – 76. [↑](#footnote-ref-32)
33. Cf. McDowell (1995), §4. Cf. also the comments about 'the cosmic exile' in his (1981), p. 248. [↑](#footnote-ref-33)
34. Cf. in this connection the sentence straddling pp. 186 and 187 in [B.] Williams (1995c). [↑](#footnote-ref-34)
35. There are indirect connections here with what Korsgaard argues in her (1986), §VI. [↑](#footnote-ref-35)
36. [B.] Williams (1995a), p. 39. [↑](#footnote-ref-36)
37. But note: had I allowed the necessity of (2c) to be 'soft' in this way, then I would also have had to modify my definition of what it is for a conative transcendental argument to be 'good'. I would have had to say that a conative transcendental argument is 'good' when it can be converted into an argument that is sound by the addition of a premiss guaranteeing, not just the desirability of its input (which is all that was needed before), but the desirability of whatever makes the 'soft' modality 'soft'. [↑](#footnote-ref-37)
38. There is no obvious reason, incidentally, why there could not be good conative transcendental arguments in which the modality *was* fairly coarse-grained, conditional on facts about our biology say: cf. in this connection Thomson (1987), passim. I take it, however, that philosophical interest in conative transcendental arguments will tend to diminish as the grain gets coarser. (Certainly this is true in the case of (ordinary) transcendental arguments, although, as I observed at the outset, one of the objections to which any such argument is vulnerable is precisely that there is no satisfactory account of the modality, conceptual and psychological elements having got irreparably tangled up.) [↑](#footnote-ref-38)
39. Cf. [B.] Williams (1995c), n. 3. This connects with a point I made earlier, in parenthesis: it is non-trivial who 'we' are supposed to be in any given conative transcendental argument. [↑](#footnote-ref-39)
40. What is required in their case, I believe, is diagnosis. See my (1997), Ch. 11, §5. [↑](#footnote-ref-40)
41. This second caveat is also meant to cover the case where the debate between internalism and externalism is a debate about explanatory priority. Thus internalists and externalists may agree that we 'must' have certain desires and that we 'must' have certain corresponding reasons, in some more or less fine-grained sense, but disagree about which of these facts explains the other. This, I suggest, can be construed as a disagreement about whether we would still have the reasons even if—*per impossibile*, as we would naturally say—we did not have the desires. In other words, it can be construed as a disagreement about finer-grained possibilities. [↑](#footnote-ref-41)
42. See esp. Kant (1933), Pt II, Ch. II, §2. See also Kant (1964), pp. 61 – 62; Kant (1956), Pt I, Bk II, Ch. II; and Kant (1978), §87. For an interesting if faint echo of these ideas in Wittgenstein, see his (1961), 6.422. [↑](#footnote-ref-42)
43. Bentham (1962), Ch. 1. And see again Harrison (1976), esp. pp. 25—28. Cf. the discussion above concerning Mill. [↑](#footnote-ref-43)
44. [B.] Williams (2006i), Ch. 4. [↑](#footnote-ref-44)
45. Wiggins (1995), all quotations from pp. 326 – 328, emphasis in original. [↑](#footnote-ref-45)
46. Korsgaard (1996), §§3.4.9 – 3.4.10. [↑](#footnote-ref-46)
47. The first premiss would then satisfy the condition that the first premiss of a non-conative transcendental argument must satisfy. [↑](#footnote-ref-47)
48. See Essay 10 in this volume, esp. §3; and my (1997), Ch. 11. [↑](#footnote-ref-48)
49. For assistance of various kinds I am greatly indebted to participants at the conference in Sheffield at which I first presented this essay, esp. to John Skorupski. I should also like to thank Bill Brewer, Robert Frazier, Penelope Mackie, Derek Parfit, and Bernard Williams for their helpful comments on an earlier draft of the essay, and Robert Stern for inviting me to participate in the conference. [↑](#footnote-ref-49)