

On the Incompatibility of Reasons Internalism and the Practical Rationality of Moral Action

Lane DesAutels

Abstract: In what follows, I explore the relationship between two widely held theses in moral and practical philosophy: reasons internalism and the practical rationality of moral action. The goal of this paper is to show the impossibility of holding a (certain sort of) internalist view of reasons while at the same time maintaining that moral acts are always necessitated by practical rationality.

§1. Introduction

The goal of this paper is to explore the relationship between two attractive and commonly held theses in practical philosophy. They are:

1. *Reasons internalism (RI)*: reasons *always* bear some relation to an agent's motivations (or desires).

And

2. The *practical rationality of moral action thesis (PRMA)*: that an act is morally right (or wrong) necessarily gives all agents to which morality applies a preemptive or decisive reason to so act (or not to act)—such that refraining from so acting would be irrational.

My aim is to argue for the following:

Incompatibility Thesis: (1) and (2) are incompatible; they cannot both be upheld without saying something absurd about morality.

Here is my plan: In §2 and §3, I distinguish between some of the varieties of reasons internalism as well as the practical force of morality thesis—briefly noting some historical and contemporary defenders of each. In §4, I give a novel argument, which I call *the Argument from Justice*, for why reasons internalism and the practical force of morality thesis cannot both be upheld.

§2. Reasons Internalisms

Reasons internalism comes in a myriad of forms. A complete taxonomy of the varieties of reasons internalism and its various associated theses is beyond the purview of this paper. Instead, I focus on (what I take to be) certain of the primary distinguishing features between kinds of reasons internalism.

Reasons internalists agree that every reason for action must bear some relation to the agent's desires or motivations. Internalists disagree, however, as to whether what is required of a consideration for it to count as a reason for an agent is some fact about an agent's actual

motivations or some other psychological state (e.g., desire) which merely plays a role in motivation. Following Finlay and Schroeder (2008), I call the former: *motivation-based reasons internalism* and the latter: *state-based reasons internalism*. To see the difference, it is instructive to note instances where the two could conceivably come apart. Suppose I desire a cookie—but lack the motivation to get up and walk to the cookie jar. Do I have a reason get a cookie? If your answer is yes, then you are a state-based reasons internalist; my desire for a cookie (despite lack of motivation to take the steps into the kitchen to retrieve one) is sufficient to give me a reason. Suppose I am motivated to make a healthy dinner of kale and beans—but have no desire to do so. Do I have a reason to cook kale and beans? If you think so, you are a motivation-based reasons internalist; my motivation to eat healthfully (despite my lack of desire to do so) is sufficient to instill me with a reason to so act.

Reasons internalists also disagree as to what sort of *relation* there must be between an agent's reasons for action and her desires or motivations. On the one hand, an internalist might require that for an agent to have a reason to ϕ , she must *actually* be motivated (or have a desire) to ϕ . On the other hand, an internalist might merely require that for an agent to have a reason to ϕ , it must be true that she *would* be motivated (or have some desire) to ϕ given some further kinds of circumstances. Call the former sort of position *actualist reasons internalism*, and the latter: *counterfactualist reasons internalism*. To complicate things further, Counterfactualists differ as to what sorts of conditions they consider to be the relevant counterfactual circumstances for a consideration to count as a reason for an agent. A counterfactualist might think, for example, a consideration is a reason for an agent to ϕ just in case (i) she has all the relevant information (and no relevant false beliefs) (Smith 1994, Joyce 2001), or (ii) she has reached a state of Rawlsian reflective equilibrium (Brandt 1979), or (iii) she has deliberated without fault from her occurrent motivations (Williams 1979) or (iv) she is practically rational (Korsgaard 1986), or (v) she was raised properly to be an ideal phronimos (McDowell 1995).¹

Due to considerations that will become clear in the following sections, the sort of reasons internalism I focus my discussion on is an actualist, state-based reasons internalism. The archetypical example of such an account is the view commonly attributed to Hume²—henceforth the *Humean Theory of Reasons* (HTR). HTR states: if an agent has a reason to ϕ , then she must

¹ Credit for the compilation of this list goes to Finlay & Schroeder 2008.

² *A Treatise of Human Nature*, Book II, Part II, Section III

have some actual desire that gets served by ϕ -ing.³ From here on out, HTR is the position I refer to by ‘reasons internalism’ (RI).

§3. The Practical Rationality of Moral Action

It would be nice if what we *morally ought* to do were to always coincide perfectly with *what it makes most sense to do*: with *what we have the best reasons to do*. If this were the case, there would never be any tricky conflicts between choosing between the moral act and the otherwise rational act.

On the face of it, however, examples seem to abound of clear instances where morally virtuous acts come at a cost to our own well-being and would, thereby, seem to be irrational for us to undertake. For example, I would be financially better off if I did not donate a percentage of my income to charity. I would have more time to myself if I were not obligated to look after my children. I could further my professional goals more quickly by lying to or taking advantage of others. Each of these is a *prima facie* instance where the morally virtuous act (acting charitably, benevolently, honestly, etc) appears to conflict with what would make life go best for me.

Many moral philosophers have gone to great lengths to construct moral systems to overcome (or otherwise make go away) this apparent conflict between the moral and the otherwise practically rational actions we might choose to undertake.

The moral philosophy of Immanuel Kant is a classic example. On Kant’s view, there is a single principle from which all specific moral duties are derived: the *categorical imperative* (CI). Kant formulates the CI in a few different ways, but argues (somewhat dubiously) that all formulations are derivable from the *formulation of the universal law*: “Act only in accordance with that maxim through which you can at the same time will that it become a universal law.” (G 4:421). For Kant, the CI is the supreme principle of practical reason. Even our hypothetical imperatives are subordinate to the CI.⁴ Furthermore, Kant argues that since morality is in the realm of metaphysics, and all reasoning about metaphysics is *a priori*, the CI must be derivable using pure, *a priori* reasoning alone. What is important about Kant for our present inquiry is this: on Kant’s view, there is no separating the moral act from the practically rational one. The two simply do not (and cannot) come apart.

³ It is worth noting, however, that my arguments may also apply to various versions of the counterfactualist account.

⁴ See Korsgaard (1997) for a detailed discussion of this.

A contemporary (non-Kantian) example of a moral philosopher who is committed to a version of the practical rationality of moral action is Alan Gibbard. Unlike Kant, however, Gibbard is firmly rooted in the contemporary metaethical expressivism movement. Expressivism about moral language is the view, roughly, that moral statements are mere evincings of non-cognitive attitudes rather than any sort of truth-evaluable claims. In his 2003 book, *Thinking How to Live*, Gibbard writes, “*Thinking what I ought to do is thinking what to do*” (x), and later:

...ought questions and reason questions are by their very nature questions of what to do. Understanding this is the way to understand what ‘ought’ assertions mean. If the chooser don’t face two clear, distinct questions, the question what to do and the question what I *ought* to do. (9-10).

On Gibbard’s account, there is no difference between thinking about what we ought (morally) to do and the question of what to do (in general). ‘Ought’ statements do not have a special property. They are simply a species of the general question of what we have to most reason to plan to do.

Derek Parfit is another recent defender of a version of (PRMA). In his 2011 book *On What Matters*, Parfit’s argues that, if morality is to matter at all, we must always have the most reason to act morally. He, like Gibbard, distinguishes between two questions: “Q1: What do I have the most reason to do? And Q2: What ought I morally to do?” He then writes,

Of these questions, it is the question about reasons that is wider, and more fundamental. And if these questions often had conflicting answers, because we often had decisive reasons to act wrongly, that would undermine morality. For morality to matter, we must have reasons to care about morality, and to avoid acting wrongly. (147)

He then goes on to defend the following position (which he labels ‘*Q*’); “When some act would be [morally] wrong, this fact always gives us decisive⁵ reason not to do it” (173). By virtue of the fact that Parfit’s *Q* only holds that the *wrongness* of an act gives us decisive reason not to do it, it might be thought of as a weaker version of (PMRA) than those formulations which claim that we always have decisive reason to act rightly. After all, unless one accepts the controversial Platonic doctrine of the unity of the virtues, there seem to be plenty of cases where different moral virtues would call for disparate courses of action. If we restrict ourselves to a version of (PRMA) like Parfit’s, on the other hand, we need not deal with having decisive reasons for acting in

⁵ Parfit defines *decisive reasons*: “When our reasons to do something are stronger than our reasons to do any other act, these reasons are *decisive*, and this act is what we have *most reason to do*; it is what we *should, ought to, or must do*” (33).

accordance with countervailing virtues like (courage and wisdom) because we only have decisive reasons not to act wrongly. For this reason, *Q* seems more easily defensible. I shall, therefore, take (PRMA) to refer to Parfit's *Q* for the remainder of the paper.

§4. An Argument for why (RI) and (PRMA) cannot both be upheld.

The general point I wish to make in this section is this: (RI) and (PRMA), in their refined versions outlined above, cannot both be upheld without running into trouble. To do this, I take the subset of moral duties stemming from justice as a starting point. Why focus on justice? I focus on justice because, as we shall see, it is the virtue that most straight-forwardly seems to require us to act in ways contrary to our actual desires. As such, I call the following argument: *the Argument from Justice*. In schematic form, it goes like this:

- (1) If (RI) holds, we only have reasons to act when such reasons connect to our actual desires.
- (2) If (PRMA) holds, all agents (to which morality applies) have decisive reason not to act unjustly—such that acting unjustly would be irrational.
- (3) There are plausible cases where avoiding acting unjustly requires of an agent to act contrary to anything in her desire set.
- (4) Given (1) and (3), if (RI) is true we *do not* always have a reason to refrain from acting unjustly.
- (5) But given (2), if (PRMA) is true we *do* always have a reason to refrain from acting unjustly.
- (6) Contradiction!
- (7) Therefore, (IR) and (PRMA) cannot both be upheld.

Premises (1) and (2) are just the respective definitions of (RI) and (PRMA) that we stipulated in §2 and §3. Premises (3)-(6), on the other hand, require some further explanation and justification.

Before offering support for the forgoing argument from justice, however, we first need to get straight as to what (precisely) *justice* is. Plato conceived of justice in the *Republic* as the overarching virtue of individuals and societies.⁶ As such, justice for Plato encompassed much (if not all) of what we now think of as morality. In today's society, however, usage of the word

⁶ Plato, *Republic*, translated and with intro by R.E. Allen, New Haven: Yale University Press, 2006.

‘justice’ has come to take a more limited meaning; it refers to only a subset of morality. As Michael Slote puts it in his 2010 essay “Justice as a Virtue”,

What individual justice most naturally refers to are moral issues having to do with goods or property. It is, we say, unjust for someone to steal from people or not to give them what he owes them, and it is also unjust if someone called upon to distribute something good (or bad or both) among members of a group uses an arbitrary or unjustified basis for making the distribution (this last aspect of individual justice obviously has reference to social or at least group justice). (1)

By way of defending premise (3) of my argument from justice (that there are plausible cases where justice requires an agent to act contrary to anything in her desire set) I take three examples from the above quotation from Slote—of paradigmatic actions prohibited by justice—and show why it is easy to conceive of an agent with no desire to so act. Having done so, I will have shown that a defender of both (RI) and (PRMA) is forced into a pernicious contradiction.

(i) Stealing: A college student desperately wants to impress her friends with the latest fashion in blue jeans, but for her, designer jeans are entirely unaffordable. She remembers having been told many times by her mother and father that stealing is wrong, but she finds herself utterly devoid of any desire to abide by the advice of her parents. The student furtively places the blue jeans in her bag and carries them out of the store unnoticed.

(ii) Not giving people what is owed to them: Imagine, for instance, that a mega-rich family member gives you a small loan so that you could make a down-payment on that new condo you, your wife, and your young children have long wanted. But further suppose that this bourgeois family member, being a bit older and busy with his retired life of golf and contract bridge, forgets all about the loan. You know that the just act would be to repay him by the time you both agreed upon, but you never liked the guy anyway—and feel zero desire to comply. He won’t even notice anyway.

(iii) Non-arbitrary distribution of goods: Consider a local politician who comes to discover an error in the state budget such that there is now a surplus of funding for the coming fiscal year. This politician knows that the virtue of justice would demand him to allocate this surplus money to road maintenance, new computers for local schools, more library books, or some other public good. However he finds himself with no desire to do so, and instead decides to keep the whole thing under wraps—quietly allocating the money to wealthy special interest groups whose support will greatly improve his chances for reelection.

What are these examples meant to show? How are they meant to support the mutual exclusivity of (RI) and (PRMA)? The point of these examples is to show plausible instances where justice requires an agent not to act in a certain way—but the agent in question has no desire to refrain from so acting. In short, these examples are supposed to show that sometimes

agents have no desire to do what justice requires. Once it is established that cases of the sort described above are easy to imagine, the rest of the argument from justice follows. Since (RI) requires of an agent's reasons that they be connected to her desires, a reasons internalist must conclude that these imaginary agents *do not have reason to refrain from acting unjustly*. But (PRMA) claims that we *always* have decisive reason not to act wrongly. And since unjust acts are certainly a species of wrong acts, we always have decisive reason not to act unjustly. In short, a defender of (RI) says these agents do not have reasons to refrain from acting unjustly; and a defender of (PRMA) says of these agents that they must have such reasons. Thus, these are plausible examples that show the incompatibility of (RI) and (PRMA).

The upshot of my argument from justice is this: anyone who wishes to be a reasons internalist and support the practical rationality of moral action runs into a problem. Namely, such a person must say of these plausible cases that the agents in the examples described above necessarily *do* and *do not* have reasons to act justly. In other words, any defender of both (RI) and (PRMA) faces a vicious contradiction.

If my argument from justice is sound, therefore, it shows that reasons internalism (of the Humean sort) and the practical rationality of moral action (of the Parfitian sort) cannot both be upheld.

§5. Conclusion

In this paper, I have defended the incompatibility of (a certain sort of) reasons internalism and the view that moral acts are always necessitated by practical rationality. To do this, I first defined and distinguished between types of reasons internalism and the practical rationality of moral action theses. I then offered an argument from justice to show that there are plausible cases which lead defenders of both (RI) and (PRMA) to reach contradictory conclusions. I conclude, therefore, that (RI) and (PRMA) cannot both be upheld and a choice between the two positions must be made.

References:

- Brandt, R. (1979) *A Theory of the Good and the Right*. New York: Clarendon Press.
- Finlay, S. & Schroeder, M (2008) "Reasons for Action: Internal vs. External" <http://plato.stanford.edu/entries/reasons-internal-external/>
- Gibbard, A. (2003). *Thinking How to Live*, Cambridge: Harvard University Press.
- Hume, D. (1739), *A Treatise of Human Nature*, edited by L. Selby-Bigge, 2nd revised edition by P. H. Nidditch, Oxford: Clarendon Press, 1975.
- Joyce, R. (2001) *The Myth of Morality*, Cambridge University Press.
- Kant, I. (1785) *Groundwork for the Metaphysics of Morals*, translated by T. Abbott, revised and edited by L. Denis, Peterborough: Broadview Press, 2005.
- Kant, I. (2002) *Theoretical Philosophy after 1781*, translated/edited by H. Allison, M. Friedman, G. Hatfield and P. Heath, Cambridge: Cambridge University Press.
- Korsgaard, C. (1986) "Skepticism about Practical Reason." *Journal of Philosophy*, 83: 5-25.
- Korsgaard, C. (1997) "The Normativity of Instrumental Reason," in *Ethics and Practical Reason*, G.Cullity & B.Gaut (eds.), Oxford: Oxford University Press, pp. 215–254.
- McDowell, J. (1995) "Might There Be External Reasons?" In J.E.J. Altham & R. Harrison (eds.), *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press: 68-85.
- O'Neill, O. (1989) *Constructions of Reason*, Cambridge: Cambridge University Press.
- Parfit, D. (2011) *On What Matters*. Oxford University Press.
- Plato, *Republic*, translated and with intro by R.E. Allen, New Haven: Yale University Press, 2006.
- Raz, J. (ed.), (1990) *Authority*, New York: New York University Press.
- Raz, J. (1975) *Practical Reason and Norms*, Oxford University Press.
- Slote, M. (2010) "Justice as a Virtue" <http://plato.stanford.edu/entries/justice-virtue/>
- Smith, M. (1994) *The Moral Problem*. Blackwell.
- Williams, B. (1979) "Internal and External Reasons." Reprinted in *Moral Luck*. Cambridge: Cambridge University Press, 1981: 101-13.