

Reply to “On the validity of a simple argument for moral error theory”
Richard Joyce

Penultimate draft of paper in *International Journal of Philosophical Studies* 24 (2016) 518-522.

I thank Kasper Christensen for his astute critical paper, which succeeds in revealing some infelicities and clumsiness in how I summed up a central argument in *The Myth of Morality* (2001). The master argument itself stretches over several chapters of that book, and though the argument no doubt has its fair share of problems, Christensen has given me no new grounds for revising its substance. But on page 42 of my book I made an attempt to state the argument in a crisp argumentative form (what Christensen calls “ARG”), and then went further and tried to present its structure in the simplest propositional form I could manage (what Christensen calls “fARG”), and I now see that these attempts at clarifying matters for the reader had some problems.

First of all, propositional logic was always going to be insufficient to capture the reasoning of ARG due to the fact that it is insufficient to properly capture the modal notion of “X regardless of Y” which lies at the heart of ARG. Christensen is quite correct that one needs to step up to modal logic in order to accomplish this. So fARG is a waste of time; I’ll happily commit it to the flames. Let us, then, examine ARG more closely.

Premise 1 of ARG is as follows:

- 1) If x morally ought to ϕ , then x ought to ϕ regardless of whether he cares to, regardless of whether ϕ ing satisfies any of his desires or furthers his interests.

When Christensen’s critique forced me to reflect carefully on this premise, I found myself puzzling over the fact that my earlier self had left the second “ought”—the “ought” in the consequent—unqualified. Surely my earlier self had meant “*morally* ought”? I even dug out the earliest draft of this chapter I could find (on a floppy disk, no less!), wondering whether a typo had slipped through. But, no, it appears that it was “ought,” not “morally ought,” all along. The version of the premise containing “morally ought” Christensen calls “(m1).” My claim is simply that I intended (m1) rather than (1) all along. I suppose the reason I left it unqualified is that it didn’t occur to me that anyone would suspect me of meaning any other kind of “ought.” I didn’t expect to be interpreted as saying “If x morally ought to ϕ , then x *prudentially* ought to ϕ regardless...” or “If x morally ought to ϕ , then x *legally* ought to ϕ regardless...,” and so on. It was pretty obviously intended as a *moral* “ought.”

What (m1) is intended to capture is the categorical nature of moral normativity. Philippa Foot (1972) famously argues that various kinds of non-moral norms seem similarly categorical. A street sign that says “No U-turn” expresses a categorical imperative; there is no tacit addendum like “...if this suits your desires” or “...if you want to obey the road rules.” Foot says the same of the norms of etiquette and of the rules of institutions such as a club. She cogently encourages us to see categorical imperatives as ubiquitous and unexceptional. In light of this, many philosophers (including myself) have seen that categoricity alone cannot account for the *authority* with which moral norms are apparently imbued. A rule of etiquette, after all, can be arbitrary and inconsequential—such as decreeing in which direction one must scoop one’s soup. A club can invent any old whimsical rule, declare that it applies irrespective of the members’

desires, and, bam, there's a categorical imperative. Surely moral normativity is supposed to have more practical oomph, more to-be-pursuedness (to use John Mackie's phrase), than this?

Some philosophers, like Kant, not only attempt to describe this additional authority (in Kant's case: that violating a moral norm is irrational) but try to show that some norms actually possess this authority. Others, like Mackie, endeavor to describe this authority (in Mackie's case: that the categoricity of moral norms is objective) as a way of showing that no norms possess it. The arguments of *The Myth of Morality* are very much along the latter lines. Pinching a phrase from Foot (1972: 309), I suggest that the extra authority of morality is that it is imbued with an "automatic reason-giving force." This gives us premise 2:

2) If x morally ought to ϕ , then x has a reason for ϕ ing.

It should be stressed that (2) really has the force of a *hypothesis* in the argument. I'm essentially saying, "Well, morality certainly seems to have *some* kind of authority that mere categoricity is insufficient to explain; let's try this." Thus if (2) turns out to be false, the master argument doesn't collapse; rather, the error theorist just returns with the question "Okay, so of what *does* the extra authority of morality consist?" After all, if it turns out that moral norms are essentially imbued with a kind of authority that no one can even articulate (let alone defend), then this is just more grist for the error theorist's mill. Of course, one might, like Foot, declare that there is really no extra authority at all—that there is nothing to be explained. And at that point the debate would head off in another direction.

My 2001 argument combines the first two premises to give (3):

3) Therefore, if x morally ought to ϕ , then x has a reason for ϕ ing regardless of whether ϕ ing serves his desires or furthers his interests.

And it is this step to which Christensen objects. But most of his efforts are expended criticizing the version of the argument containing premise (1) rather than (m1). Everything he says in criticism of that argument is correct; I just wish I could have saved him the trouble by hopping in a time machine, going back to the moment at which I was about to submit my manuscript to CUP, and inserting the word "morally." The addition of this word makes not one iota of difference to anything else said in my book, and yields an argument that Christensen admits is valid.

But Christensen is still not entirely happy with the valid form. Toward the end of his paper he expresses concern that "(m1), as compared to (1), is too strong to convince anyone who is not already convinced of something like moral error theory" (#). And he worries whether the evidence that I marshal in support of the (1) will also support (m1). I think both these concerns are unfounded.

First, recall that (m1) is simply an attempt to state the categoricity of moral norms, which is something that many moral realists want to uphold. Kant himself wouldn't hesitate to endorse (m1)—indeed, it is a cornerstone of his moral philosophy—yet it is hardly Kant's sympathies with a moral error theory that lead him to this! So the claim that (m1) is likely to be attractive only to those antecedently sympathetic to moral skepticism is false.

Second, given that my stating the first premise as (1) rather than (m1) was an oversight rather than a deliberate decision—given, that is, that I intended readers to interpret the premise as (m1) all along—the evidence I marshaled *was* in support of (m1). What Christensen evidently needs to

do is read over the sections of my book that attempt to establish this premise and demonstrate that my arguments were aimed at establishing (1) rather than (m1). But he doesn't do this, and if he attempted to it would quickly become apparent that my arguments were always focused on (m1).

How does one establish that a kind of imperative is categorical rather than hypothetical? A hypothetical imperative is essentially a piece of advice on how to satisfy one's ends. If the speaker mentions the ends in question, or is willing to make them explicit rather than tacit upon questioning, then the imperative is probably hypothetical. If it is discovered that the subject lacks ends which it was previously thought s/he has, and this leads to an imperative being withdrawn, then it is probably hypothetical. For example:

A: "You ought to hurry if you're going to catch the 2:30 bus."

B: "But I've changed my schedule; I'm going to catch a later bus."

A: "Okay, then I guess you don't need to hurry."

In *The Myth of Morality* I attempted to persuade readers that moral imperatives are not like this—they are not withdrawn when it is discovered that the transgressor lacks ends that would be served by acting in accordance with morality. Compare:

A: "You ought not punch babies for fun."

B: "But I *love* punching babies, and I can avoid all the negative consequences." [Here B produces the Ring of Gyges.]

A: "Okay, then I guess it's all right for you to punch babies for fun."

The weirdness of this second exchange (and others like it), in comparison to the unexceptional first exchange, is the kind of evidence I raised in support of the first premise. (And note again that it is not merely those already sympathetic to moral skepticism that are supposed to be persuaded.) But all my examples pertained to *moral* "ought"s. Even though the "ought" uttered by A in the second exchange is unqualified, no one who reads the relevant sections of my 2001 book could be in any doubt that I'm discussing a special feature of *moral* normativity.

Here's an analogy. Suppose I wanted to point out some special geometric feature of scalene triangles. A simple way of stating it would be this:

i) Scalene triangles have property P.

But for various reasons, I might want to express this as a conditional:

ii) If x is a scalene triangle, then x has property P.

Or I might want to introduce an agent's standing in some relation to scalene triangles:

iii) If x is drawing a scalene triangle, then x is drawing something with property P.

Or I could put it this way:

iv) If x is drawing a scalene triangle, then x is drawing a triangle with property P.

Most of the trouble raised by Christensen boils down to the fact that when I wanted to call attention to the categoricity of moral norms—the equivalent of (i)—I expressed things in the form of (iv). But it should be quite clear that the second appearance of “triangle” in (iv) can be explicitly restricted to *scalene* triangles:

v) If x is drawing a scalene triangle, then x is drawing a scalene triangle with property P.

The analogy of Christensen’s criticism would be that I have confused (iv) and (v), and gathered evidence for (iv) when I really need to be putting forward evidence for (v). But that criticism would be misplaced, for the evidence for (i), (ii), (iii), (iv), and (v) *will all be the same*. In the metaethical argument, what I am really endeavoring to establish is the equivalent of (i)—I’m simply aiming to show that moral norms are usually uttered with categorical force. Whatever may be weak or unconvincing in the arguments I offered in support of this proposition, it is not because I was mistakenly endeavoring to establish the wrong premise. Christensen ends his paper by challenging me to state succinctly what the valid argument for the moral error theory is supposed to be. The answer is straightforward: it is the argument with (m1) as the first premise. And the retort that *The Myth of Morality* devotes energy to establishing (1) rather than (m1) is simply mistaken.

References

- Christensen, K. Forthcoming. “On the validity of a simple argument for moral error theory.” *International Journal of Philosophical Studies* #.
- Foot, P. 1972. “Morality as a system of hypothetical imperatives.” *Philosophical Studies* 81: 305-316.
- Joyce, R. 2001. *The Myth of Morality*. Cambridge University Press.
- Mackie, J. 1977. *Ethics: Inventing Right and Wrong*. Penguin.