Kant, Wood and Moral Arguments

Andrew Chignell

Princeton University, Princeton, NJ, USA
Email: chignell@princeton.edu

Abstract

In this article I discuss the moral-coherence reading of Kant’s moral argument offered by Allen Wood in his recent book Kant and Religion, display some of the challenges that it faces and suggest that a moral-psychological formulation is preferable.

Keywords: moral arguments; God; the soul; moral psychology

1. Introduction

Kantian moral arguments start from some sort of justified commitment regarding the ‘practical’ (i.e. what we ought to do) and end with a justified commitment to something ‘theoretical’ (i.e. what exists). There are variations: in some, the starting commitment takes the form of mere hope; others start with a stronger attitude like knowledge. The results also vary: some moral arguments deliver rational faith; in others, the result is rational hope. But again, the essential feature is the move from ought to is – from some kind of justified attitude about moral considerations to some kind of justified attitude regarding what exists.

In this brief article I consider two varieties of moral argument in Kant – I will call them moral-coherence arguments and moral-psychological arguments.

The moral-coherence variety says that rational coherence in our practical lives presupposes certain theoretical commitments. Even if the ‘crooked timber of humanity’ never produces something straight, our rational nature demands that we set the Highest Good (conceived as happiness for each person in precise proportion to their degree of virtue) as our ideal ultimate end. Reason also then justifies us in committing to the existence of whatever is required to explain the real possibility of the Highest Good. We find this variety featured prominently in Kant’s second Critique (1788) – it is what people typically think of when they refer to ‘Kant’s moral argument’.

The moral-psychological variety, by contrast, says that shoring up and sustaining our resolve with respect to our moral commitments requires (for many of us, much of the time) certain theoretical commitments, and that the latter can be prima facie morally justified in virtue of that resolve-sustaining ability. We find the moral-psychological variety articulated most prominently in Kant’s third Critique (1790) and Religion (1793).

In Kant and Religion (2020) Allen Wood champions a moral-coherence argument for a commitment to the existence of God and a Future Life (call this conjunction ‘GFL’). He also rejects moral-psychological arguments as irrational. Wood is writing in

© The Author(s), 2022. Published by Cambridge University Press on behalf of Kantian Review.
a historical mode, trying to reconstruct Kant’s own efforts on behalf of rational Belief (Vernunftglaube) in GFL. Mark Johnston has recently developed a moral-coherence argument in a constructive, contemporary mode; I discuss aspects of it in section 3, though a full discussion of Johnston’s argument would be a project for another day.

My goals here are to lay out the moral-coherence argument offered by Wood, display some of the challenges that it faces, and suggest that a moral-psychological formulation is preferable, despite concerns about irrationality expressed by both Wood and Johnston.

2. Wood and Kantian ‘Belief’

In chapter 2 of Kant and Religion (2020), Allen Wood returns to the topic of Kantian moral arguments – arguments that his younger self did a great deal to revive. The early Wood (1970) focused on the moral-coherence variety of the proof that we find in the second Critique and some of the lectures. In Kant and Religion, recent Wood takes us carefully through the three Critiques and discusses the presentations of the moral proof in each, arguing that they are all best read in the moral-coherence way. In this section I look at Wood’s characterizations of the proof in the first and second Critiques. In the final section I consider his criticisms of the moral-psychological reading.

2.1 Moral argument in the first Critique

The proof as it is presented in the Canon of Pure Reason of the first Critique appears to involve the oddly un-Kantian insistence that hope for happiness in a future life is required for us to be morally motivated:

Thus without a God and a world that is not now visible to us but is hoped for, the majestic ideas of morality are, to be sure, objects of approbation and admiration but not incentives for resolve and realization (Gegenstände des Beifalls und der Bewunderung, aber nicht Triebfedern des Vorsatzes und der Ausübung) because they would not fulfill the whole end that is natural for every rational being and determined a priori and necessarily through the very same pure reason. (A812/B840, italics added)

Not just hope but also fear: Kant mentions both positive and negative consequences as part of our moral motivation.

Everyone also regards the moral laws as commands, which, however, they could not be if they did not connect appropriate consequences with their rule a priori, and thus carry with them promises and threats (Verheissungen und Drohungen). (A811/B839, bold in original, italics added)

On its face, this suggests that the moral law is only normative for us (only a ‘command’) if it is connected to promises and threats. Thus these texts depict us as morally motivated not by respect for the moral law alone, but also by the promise of glory and the fear of perdition.
Because this sounds so un-Kantian at first face, some authors argue that the Canon contains merely a preliminary sketch of Kant’s argument, written at a time when he had not yet worked out his ethics very clearly. Thus Henry Allison says that the motivation Kant invokes in this passage is at least partly eudaimonistic and heteronomous – it requires an appeal both to what the moral law commands and to our desire for happiness. Kant later abandons this picture, says Allison, in favour of his classic view on which the moral law is the only appropriate and also fully sufficient rational motivation.⁵

Wood, by contrast, argues that these passages from the first Critique are written in the same spirit as Kant’s mature moral psychology. The claim in the Canon, after all, is not that we ought to will (the Highest Good) is just a state of happiness (as in standard eudaimonism). Rather, it is happiness conditional on and apportioned to moral worth. ‘This means that his position is not eudaimonistic’, says Wood, ‘if eudaimonism is the position that morality as a whole is to be derived from the end of happiness’ (p. 40, my italics). Wood also argues that Kant is not endorsing what he would later refer to as ‘heteronomous’ motivation. Human beings are essentially a combination of both rationality and sensibility, says Wood, and so the Highest Good that creatures like us would autonomously pursue is a complex state that appeals to both aspects of our being. (Wood admits that ‘autonomous’ incentives in later works would exclude appeals to sensibility, however (p. 40.)

This seems right thus far. But setting terms like ‘eudaimonism’ and ‘heteronomy’ aside, there is something to Allison’s point nonetheless. Key to Wood’s reading of the Canon is a distinction between our ‘rational incentive for moral conduct’ and the ‘rational incentive we have to do what we know we are obligated to do’ (p. 42). The first, he says, is simply a recognition that an action is commanded by the moral law (and thus, derivatively, by God as well). He quotes Kant here:

So far as practical reason has a right to lead us, we will not hold actions to be obligatory because they are commands, but rather regard them as divine commands because we are internally obligated to them. (A819/B847)

Wood takes this to show that, for Kant in the first Critique, ‘we are aware of an unconditional obligation to obey the moral law, grounded on practical reason alone’ (p. 42). But that rational incentive (for moral conduct) is supposed to be different from the rational incentive to do what we know we ought to do. It is with respect to the latter, Wood suggests, that we must appeal to ‘the rational hope that we will enjoy the happiness of which we are thereby made worthy’ (p. 40).

I find the distinction hard to understand. If we have a ‘rational incentive for moral conduct’, then it seems like we ipso facto have a rational incentive to engage in that conduct – i.e. to ‘do what we know we are obligated to do’. Indeed, it seems like these are almost two ways of saying the same thing. So if the rational hope for the degree of deserved happiness is required for us to have a rational incentive to do what we ought to do, then it looks like it is also part of the rational incentive for moral conduct.

That is also what the texts suggest. Kant says in the passage quoted above (A811/B839) that the commands of the moral law ‘could not be’ commands if they did not carry with them various post-mortem promises and threats. But if a principle could not be a command without a connection to ‘appropriate consequences’, then it is hard
to see how it could constitute a ‘rational incentive for moral conduct’ all on its own. Wood wants to say that we can regard it as a command (obligation) and thus have a rational incentive to act, but at the same time lack a rational incentive to go ahead and perform the action. I don’t see much daylight between the two concepts.

But even granting the distinction, it remains unclear how the hope for happiness in proportion to our worthiness counts as a ‘rational incentive to do what we know we morally ought to do’. It is true that the hope itself is rational just insofar as its object is not unconstrained happiness but rather the happiness of which we are morally worthy. That is the degree of happiness that we may (dare) rationally hope for, according to Kant. But the fact that the hope is rational in this way need not make it part of our rational incentive.

We can see this by comparing the first Critique picture to the story in the *Groundwork* and second Critique, in which (as Wood puts it) ‘pure practical reason, duty, or the moral law all by itself constitutes a sufficient rational incentive for moral action’ (p. 41). In those works, Kant still thinks that we rationally hope for deserved happiness. But the presence of such hope does not automatically make it part of our rational incentive. In other words, the fact that the hope is rational (by way of being constrained by moral worthiness) does not mean that it rationally motivates. We would need a further argument for thinking that hope (or rather its object, which is what I think Wood means here) can be a rational incentive for ‘doing what we know we ought to do’.

A related problem with Wood’s reading of the first Critique is this: we know that the Highest Good for Kant essentially involves a state in which happiness is perfectly apportioned to our degree of ‘moral worthiness to be happy’. But what is moral worthiness to be happy? For Kant it is a function not of our behaviour or its consequences, but of the extent to which we take the moral law alone to be our rational incentive for acting. But if that is right, then it is hard to see how the Highest Good as a whole could be what rationally motivates us, given that the degree to which we take the moral law alone to be our rational motive is a component of the Highest Good.

Put another way, it is unclear how both of these can be true:

1. Happiness in proportion to moral worth is part of my rational incentive

and yet

2. Moral worth is measured by the extent to which the moral law (alone, apart from any connection with happiness) is my rational incentive.

It is clear, of course, that Kant is not opposed to the idea that we can in principle be motivated by both rational and affective considerations, or that they can come in combination. We see this in the promises and threats passage (A811/B839), and in his effort to find a middle way between the Epicureans and the Stoics when developing the doctrine of the Highest Good. It is also clear that Kant believes we can take the moral law as a sufficient rational incentive for action and then hope for proportional happiness. That is typically what Kant means when he speaks of ‘willing’ or ‘promoting’ the Highest Good. But both of these claims are different from Wood’s claim that divine promises and threats are part of our rational incentive.
In the end I think we have to allow that in the first Critique the rational incentive provided by the moral law receives an essential supplement by way of an affective incentive for happiness and the avoidance of suffering (carrots, sticks!). Yes, the Highest Good appeals to both our embodied and rational natures, but it does not make happiness – even happiness apportioned to moral worth – a part of the rational incentive. I am not sure whether this is eudaemonistic or heteronomous; that depends on how we define the terms. What is important is that a non-rational incentive is playing an essential role in motivating moral action. As a result, this looks like a picture that the later Kant would condemn.

The moral-psychological argument, by contrast, distinguishes between the rational incentive for moral conduct – which consists in the recognition of the moral law alone – and an empirical-psychological need (in many of us, at least) to shore up our moral resolve. Kant’s first Critique formulation foreshadows that version of the moral argument, although it rather crassly suggests that what is required are heavenly carrots and hellish sticks. But it also erroneously (from Kant’s mature point of view) says that the moral law does not count as a command unless those incentives are in place. The moral-psychological formulation, by contrast, focuses on what we (many of us, at least) might need by way of hoped-for consequences not in order to be rationally incentivized but rather to shore up our moral resolve. I return to this in section 3.

2.2 Moral argument in the second Critique

Wood’s main focus in Kant and Religion is the classic version of the proof offered in the second Critique and related writings. This is the moral-coherence argument according to which setting the Highest Good as our end rationally commits us to the existence of whatever is required to make that end really possible. It turns out, for Kant, that what is required is not just the this-worldly efforts of individuals or collectives, but also (given the perfect justice involved) an extended and possibly everlasting future life superintended by a just deity. Thus, GFL. Wood lays it out this way:

1. I have an obligation to will the Highest Good as an end.
2. To will something as an end presupposes that it is possible.
3. Therefore, someone who wills the Highest Good as an end must assent to the conditions of its possibility.
4. The Highest Good is possible only if there is a God who apportions happiness to worthiness.
5. Therefore, if I fulfill my obligation to will the Highest Good as an end, I must assent to the existence of God. (p. 46)

There are deep and notorious problems here. The first and most obvious has to do with (1): why ought we to will the Highest Good rather than simply our own virtue? This is an item of perennial debate in the literature. Another question about (1) is why we ought to will the complete Highest Good, instead of just the promotion or approximation of it. Also well-travelled ground.

There are additional questions about (2) – in particular, about the ought-imply-can principle and whether it can deliver the sort of ‘real possibility’ needed to make (3) plausible. Johnston for instance openly wonders why we cannot coherently will an
end that we presuppose to be merely logically possible, without any commitment regarding real possibility (Johnston 2019: 51).

Finally, there are questions about why the actuality of GFL is required to ground the possibility of the Highest Good. Why is the real possibility of GFL not a sufficient ground of the real possibility of the Highest Good? There is also an under-discussed problem with (3). Even if it is true that willing something presupposes that it is possible (as in (2)), that says nothing about whether we ‘must assent’ both to its possibility and the conditions thereof. To presuppose that it is possible for me to get a beer from the fridge and then set that as my end need not involve any actual assent regarding the possibility in question, much less require one. In the context of acting, I count as presupposing that it is possible (and hoping that it is true) simply by lacking the certainty that it is impossible.

Wood sets these difficulties aside for the most part in order to focus on (4). The question here is why the existence of a full-blown classical God is required to ground the possibility of the Highest Good. Why could not one of Hume’s mediocre deities suffice, or a law of karma, or even just nature itself conceived as what Fichte calls a ‘living moral order’? Wood argues that Kant may have been trying to avoid this objection (and the threat of a ‘shrunken’ picture of the divine) by making the classical God an ‘immediately certain postulate’ that attends the willing of the Highest Good, rather than something whose characteristics must be drawn out of the moral argumentation itself. ‘Because this action is morally required’, Wood says, ‘the postulate becomes immediately certain for practical reason, even if it remains indemonstrable and uncertain for theoretical reason’ (p. 47).

It is true that Kant often refers to God and the Future Life as ‘postulates of practical reason’ in the second Critique and elsewhere (5: 122ff.). But people who are looking for an account of rational theistic Belief are unlikely to want to leap-by-postulate into a conception of God that goes far beyond anything the argument itself motivates. Wood acknowledges Kant’s efforts to show that God conceived in the classical way is required to ensure the real possibility of the Highest Good. But he sets these aside as a function of ‘the depth of [Kant’s] commitment to the traditional scholastic-rationalist concept of God’ (p. 48).

My own sense is that Kant thinks his argument that only a traditional God could guarantee the possibility of the Highest Good works. But Wood is right that it does not, at least as stated, and so on his reconstruction we are left with a postulation whose ‘immediate certainty’ is just stipulated rather than motivated. The moral-psychological formulation of the argument does better here as well.

3. The moral-psychological argument

We have looked briefly at Wood’s account of the moral-coherence variety of Kantian argument and considered its benefits and costs. We can now turn to the moral-psychological version of the argument to see whether it does any better. (Spoiler alert: I think that in many respects it does, and is thus the superior variety of Kantian moral argument.)

The moral-psychological argument, in brief, says that appeals to morally important but contingent psychological needs can justify certain commitments (including to GFL) under certain circumstances for certain subjects. If someone needs to be able
to hope that there is some sort of ultimate moral order to which their efforts contribute just in order to keep soldiering on in the moral life, then that need itself provides prima facie moral justification for Belief in GFL. I have reconstructed what I take to be Kant’s argument on this score elsewhere (Chignell 2018 and forthcoming). Here I am going to articulate it in terms introduced by Mark Johnston (2019) in his contemporary discussion of the moral argument. (Note, however, that Johnston’s own version is not of the moral-psychological but rather the moral-coherence variety.)

Let us start by saying that in order to remain fully resolved in our commitment to the moral life we must at least be able to hope that:

1. There is an Ethical Economy in which proportionality obtains, and which involves a Reckoning governed by a just cosmic Administrator (EERA).

Rational hope presupposes the real possibility of its object. So now we have moral justification – grounded in the moral-psychological need for such hope – for a commitment to:

2. EERA is really possible.

Johnston (2019: 51) criticizes other versions of the moral argument on the grounds that an ought-implies-can inference may not be able to deliver real rather than merely logical possibility. Here however we are going from the moral need to hope that p directly to moral justification for presupposing that p is really possible. In other words, to the presupposition that beings and powers in the actual world could in fact bring p about. That is the only kind of possibility that would satisfy the need and so the need is justified. But as Johnston himself points out, EERA is really possible only if

3. We are capable of Multiple Embodiments (MESS).

And the best or perhaps only possible explanation of MESS involves appeal to:

4. We are Real Enduring Substances (RES).

Thus, the argument concludes, our need to sustain moral resolve provides (defeasible) moral justification for Belief that MESS+RES, and this is sufficient to repudiate ‘reductive naturalism’ about the self.

In short, we start with the rational need to avoid morally deplorable demoralization and sustain morally significant hope, and that gives us moral justification for Belief that MESS first, and then that RES. The reductive naturalist is dispatched on practical grounds, and we avoid some of the main problems with the argument construed as a moral-coherence variety.

4. Objections and replies

We might still wonder why, among all the other possible grounds of an ultimate moral order, a full-blown classical deity with all its omni-properties is required to administer the ethical economy and the Reckoning (EERA). I am sympathetic to the objection,
obviously, since I raised an analogue of it in the discussions of moral-coherence arguments above. But there it was a matter of pointing out a conceptual fact: the classical God is not the only coherent explanation of the real or practical possibility of the Highest Good or the ethical economy.

Here the moral-psychological argument has an advantage, since it is based not merely in conceptual facts but also in empirical facts about what is required for average agents to sustain moral resolve. I suspect that some and perhaps many people would report that the existence of a moral order – even the necessary existence of a moral order – would be impressive to them but also insufficient for sustaining or stabilizing resolve. But they might also report that, as a matter of psychological fact, they will not be able to soldier on in the moral life unless they can regard this order as both intended and superintended by a being who is also intimately aware of and cares about their efforts, who arranges for their encouragement in various ways, who ensures that there are exemplars in history and scripture for them to emulate, and who shows the occasional mercy when they fail. They have a moral need, in other words, to conceive of the order as governed not by tit-for-tat impersonal laws but by a personal superintendent who does things for reasons, sets up the system from a motive of both justice and love, and recognizes individual efforts, great and small.

Even granting this, an objector might say, cannot such people find an equal amount of psychological sustenance in the thought that the world is superintended by an extremely strong, smart and well-intentioned Humean deity, even if it is not quite the omni-God of the philosophers? Here again I think the moral-psychological version of the proof offers a unique way to resist. For it may be that, as a matter of fact, many of us need to hope that the superintendent of the moral order is perfectly good, so that our dependence on that order can be complete and not provisional. We might also need to hope that the system is entirely stable and cannot even in principle be undermined by some other, even stronger power. This would then offer a moral-psychological ground for Belief in an all-powerful and perfectly good superintendent of the universe.

By now the tough-minded evidentialist objector will be apoplectic: how can it be acceptable to Immanuel Kant, reason’s bulldog, for weak-minded people to go from a mere need to hope for comic justice to full-blown Belief in the existence of the classical deity? Why is this not just rank self-deception or a pernicious form of Feuerbachian wish-fulfilment? Wood expresses this objection forcefully:

Using belief in God to supplement our weak moral motivation is an interpretation of Kant’s moral arguments that has been explored by Adams (1987, pp. 144–164; 1999, pp. 384–391) and Chignell (2018). That suggests an argument that I regard as fundamentally different from Kant’s. That the offered psychological benefit is moral may distract us from the fact that this argument is not fundamentally different from the corrupt bribe offered you to believe that seven is not a prime number. No promised advantage or benefit, even a moral one, could ever give you any reason whatsoever to assent to God’s existence. It offers an instrumental reason for doing something corrupt and self-deceptive to yourself that brings about nonrational assent. In Kant’s terms, it grounds persuasion (Überredung) rather than conviction (Überzeugung) (A821/B849). (Wood 2020: 41, n. 11)
Johnston (2019: 51) agrees. To start with hope and move from there to a ‘positive account of what we essentially are is likely to seem to be sheer metaphysical reverie’.

The concern is real: even in a context where, as Kant says, the practical has ‘primacy’, we cannot let practical needs and interests turn theology into a wax nose (this is Locke’s metaphor for a thing whose appearance is easily bent and disfigured by our every whim). Still, I think the argument as reconstructed here offers a few *principled* things to say in response.

First, Wood is right that there is an air of persuasion here, but Kant himself says that it is permissible to hold onto persuasion in cases where it is subjectively useful. ‘I can preserve persuasion for myself if I please to do so, but cannot and should not want to make it valid beyond myself’ (A822/B850). Kant is more permissive on this score, it seems, than Wood.

Second, remember that this is rational Belief not belief, so there are no violations of evidentialist norms on belief here.

Third, there may be some moral saints who can sustain their resolve even in the face of what Kant calls the ‘abyss of purposeless material chaos’. These will be few and far between – rarer even than the ‘righteous Spinoza’. But if they exist at all, then to them Kant’s argument has nothing important to offer.

Fourth, there may be others – not quite saints but moral pillars, perhaps – who can sustain their moral resolve in the face of great obstacles by adopting Belief that nature itself or a karmic law will someday bring about the Highest Good (or the EER in EERA). Their substantial hopes do not require Belief that the moral order is superintended, and Kant would simply commend them on their way.

Fifth, it is important to see that, even with respect to those for whom Belief in a suprême superintendent is morally essential, Kant is not recommending self-deception. We must keep before us the facts that (a) there is no persuasive empirical or theoretical evidence for or against the existence of such a being, (b) the Belief that the argument underwrites is *morally* (not epistemically) justified, and (c) this moral justification is defeasible to boot. If we find out later that there are moral disadvantages to this sort of Belief – that adopting it either constitutes or leads to the violation of some other duty, or fails to sustain the resolve that it promised – then it will have to be given up.

Finally, this is not wish-fulfilment of the sort that would take even the most improbable proposition to be true, just in order to feel better (or receive a bribe). It is not that because (d) we have to presuppose that the outcome is really possible, (e) the existence of the adequate ground of that real possibility (i.e. GFL) must be epistemically ambiguous or even theoretically undecidable, and (f) the goal is not to feel better or get rich but rather to receive *essential* moral sustenance. In other words, this is not just any old interest or need; it is ‘a rational interest’ (A806/B833) – ‘a need attaching to reason in itself’ (‘Orientation in Thinking’, 8: 136).

Wood is right that Kant clearly does not want us to abuse our theoretical faculties by forcing ourselves or others into full-blown beliefs about these matters. It is also true that we should not degrade ourselves or others by treating morally neutral ritualistic acts as though they were somehow pleasing to God. But those of us who are frailer than the moral saints may still need to cling to key, resolve-sustaining hopes. And thus we may also be (defeasibly) morally justified in holding the Beliefs required to ground those hopes.7
Notes

1 I use the capitalized 'Belief' here and throughout to translate Kant's Glaube, for which I think there is no suitable English translation. Wood does the same.


3 See Allison 2011: 55.

4 I have left off the parenthetical justifications for each step that Wood provides in the text, just for the sake of brevity.

5 For the suggestion that a commitment to God's real possibility is religiously sufficient when combined with hope, see Muyssens 1979; compare Wood 1991 and Chignell 2014.

6 For more on this minimal modal constraint on hope, see Chignell 2014.

7 Thanks to an anonymous referee for helpful comments on an earlier draft.

References


Cite this article: Chignell, A. (2022). Kant, Wood and Moral Arguments. Kantian Review 27, 61–70. https://doi.org/10.1017/S1369415421000716