Kate Moran

A Modest Defense of the Personal Highest Good

Abstract: In this essay, I argue that Kant has a philosophically coherent notion of the highest good for the individual, where this is understood as an object of practical reason, but not as an object of moral action. Crucially, only an object of moral action demands a causal relationship between virtue and happiness, according to which the former brings about the latter. The account I argue for here requires only a conditioning relationship between virtue and happiness. As such, the account I offer is a relatively scaled-back version of the highest good for the individual. Nevertheless, I think it can serve as the basis for a kind of practical belief.

Keywords: highest good, happiness, postulates, laws of nature

The main argument I want to make in this paper is that Kant has a philosophically coherent notion of the highest good for the individual, where this is understood as an object of practical reason, but not as an object of moral action. The highest good for the individual consists of an individual agent's maximal virtue, combined with as much happiness as is consistent with such virtue for that agent. As such, it is the object of practical reason for individual rational yet sensible agents whose reason necessarily seeks a kind of totality. This notion of the highest good is distinct from an account of the highest good according to which moral action brings about happiness generally, within a moral community and over the course of history. A crucial distinction between the two accounts, I will argue, is that only the second account demands a causal relationship between virtue and happiness, according to which the former brings about the latter. The first account only requires a conditioning relationship between virtue and happiness, according to which the former limits the how much of the latter an agent can enjoy.

It would be more accurate to say that these are the various relations between virtue and happiness that Kant should have recognized in each account of the highest good. Infamously, Kant insists on a causal relationship between the individual agent's virtue and her happiness in the Dialectic of the *Critique of Practical Reason*. This assertion, for which he gives only the suggestion of an argument, is at the foundation of much of what is confusing and problematic about the argument for the postulates of pure practical reason in that text. First and foremost, it is what leads some philosophers to wonder if Kant is slipping into heteronomy, since it introduces the

https://doi.org/10.1515/9783111544731-005

prospect of an agent's own happiness into her deliberations about moral action.[1] Beyond this, it is at the basis of Kant's arguments for the existence of God and the immorality of the soul, which, if not exactly 'train wrecks', are something like an overly-exuberant stumbling.[2] Nevertheless, there are passages in which Kant does not obviously insist on a causal relationship between individual virtue and individual happiness, and in these passages, he offers his most philosophically coherent notion of the individual highest good.

The discussion that follows proceeds in three sections. In the first, I review some of the problems that commentators, including Paul Guyer, have noted about Kant's individualist account of the highest good. I also give an overview of the account of the highest good that places its realization in the moral community, and discuss some of its interpretative virtues. Despite the strengths of the latter account, I suggest that there are independent reasons to take seriously the notion of the highest good for the individual. In the second section, I make a case for a personal account of the highest good that does not insist on a causal connection between virtue and happiness. In the final section, I explore what agents might have practical grounds to believe on the basis of this individualist account of the highest good.

A virtue, I hope, of the account I offer here is that it allows two distinct notions of the highest good to exist side by side, since each is concerned with a different set of questions. One is a set of questions about the totality of virtue and happiness for a rational and sensible agent. The other is a set of questions about whether and how we can expect virtue to bring about happiness. Kant's exuberant mistake in the *Critique of Practical Reason* is to attempt to allow the first account of the highest good to answer the questions belonging to the second.

1 Most notably, perhaps, Lewis White Beck: "Kant simply cannot have it both ways. He cannot say that the highest good is a motive for the pure will, and then say it is so under the human limitation that man must have an object which is not exclusively moral (for there is nothing moral in happiness except insofar as its condition is worthiness to be happy, and even then the moral value lies in the worthiness, not the enjoyment)." (Beck: *A Commentary on Kant's Critique of Practical Reason*, Chicago 1960, 244.) But the concern is probably overblown, since whatever happiness is part of the agent's highest good *qua* individual is still conditioned by morality. In Kantian terms, the maxim of self-love is still subordinate to the morality (RGV, AA 06: 36.21–23).

2 Cf. Paul Guyer's discussion of the third section of the *Groundwork*: "Naturalistic and Transcendental Moments in Kant's Moral Philosophy," *Inquiry* 50, 2007, 444–464, 445.

1 Two Accounts of the Highest Good

In the Dialectic of the *Critique of Practical Reason*, Kant famously characterizes the highest good as the complete good for rational and sensible agents,[3] and as the object of practical reason.[4] Importantly, the highest good is, for Kant, composed of two parts: first, perfect or maximal virtue, and second, the greatest degree of happiness that is consistent with such virtue. On this point, Kant sets himself apart from the ancients, who attempted to fuse the two into a relation of identity. For the purposes of the discussion in this essay, it is important to take note of Kant's objection to the Stoic account of the relationship between virtue and happiness. As Kant interprets that view, "consciousness of one's virtue is happiness."[5] Kant agrees with his picture of the Stoics that consciousness of one's virtue — to the extent that this is at all possible — generates a positive feeling, something that he characterizes as 'contentment'.[6] But this is not the same as happiness. Without entering into the robust debate about how best to understand the Kantian conception of happiness, we can minimally note that Kantian happiness includes the satisfaction of various needs and desires that originate from human beings' sensible nature.[7] Kant's notion of the highest good thus brings together the perfection of our rational nature (virtue) with as much satisfaction of our sensible nature (happiness) as is consistent with such virtue.

The preceding is only a rough sketch of Kant's general conception of the highest good, and the details of this conception do not always remain stable or consistent throughout the texts in which he grapples with it. In particular, Kant sometimes argues that the highest good must find its realization within the individual agent, i.e. that the highest good describes the perfect virtue of an individual, and the greatest degree of happiness that is consistent with such virtue for that individual. This is what Guyer sometimes describes as the 'composite' account of the highest good, since on such an account "the highest good [is] the compound object of a compound being, the natural pursuit of personal happiness restricted by the moral requirements to fulfill duty and achieve virtue as the condition for worthiness to be happy."[8] In the Dialectic of the *Critique of Practical Reason* this composite account

3 KpV, AA 05: 111.01.

4 KpV, AA 05: 109.20.

5 KpV, AA 05: 111.24–25.

6 KpV, AA 05: 117.28–29.

7 For a discussion of these, see Timmermann, Jens: *Kant's Will at the Crossroads.* Oxford 2022, chapter 2.

8 Guyer, Paul: "Ends of Reason and Ends of Nature." In *Kant's System of Nature and Freedom*. Oxford 2005, 169–197, 183.

of the highest good is meant to give us practical grounds for belief in several items of special metaphysics. Kant's argument for this conclusion is that the highest good, as the object of practical reason, is morally necessary. In particular, Kant argues, virtue and happiness together exist together in a synthetic relationship in which virtue brings about happiness. As he describes it, the connection between virtue and happiness "is found in virtue's producing happiness as something different from the consciousness of virtue, as a cause produces an effect."[9] Of course, this causal connection between virtue and happiness is not one that we typically see in this world, so Kant thinks we must postulate God as the author of laws of nature that conjoin virtue and happiness in a causal relation, and we must postulate the immortality of the soul, since perfect virtue — the first part of the highest good — is impossible in this life.

There are too many problems with this argumentative strategy to give an exhaustive account of them here. It makes good sense to point out that, as sensible agents, we aim at our own happiness, and that this happiness is in fact good insofar as an agent is worthy of it, i.e. insofar as that happiness is consistent with morality. But it is not at all clear what argument there is for the claim that morality demands a reward of happiness, or that an individual's virtue must bring about that individual's happiness, either in this life or the next. This is a point I will return to in more detail in the next section.

Setting that point aside for the moment, there is a further question about why Kant should insist that we are incapable of complete virtue at any moment of our existence. The most straightforward way of interpreting this claim is that we are constitutionally incapable of perfect virtue, owing to the fact that we are in the rather unfortunate position of having to drag our sensible selves around with us everywhere we go, which in turn means that we are constantly beset by temptation in the form of inclination. This would be to equate virtue with holiness, which Kant seems happy to do in the *Critique of Practical Reason*,[10] but it is not how he describes virtue in other texts. In the "Doctrine of Virtue," for example, he describes virtue as a kind of moral strength in the face of contrary inclination, not as the expulsion of inclination.[11] And, as Guyer has argued in detail, this account of moral progress or reform is not consistent with Kant's later account in the *Religion Within the Limits of Mere Reason*, where he argues instead that we are capable of a change of heart

9 KpV, AA 05: 111.16–17.

10 KpV, AA 05: 122.09.

11 TL, AA 06: 405.15–16.

in this life, though we cannot be certain of this fact since the change of heart occurs noumenally.[12]

On a similar note, if Kant's assumption is that our sensible nature is all — or even most — of what stands in the way of achieving perfect virtue, it is not clear why he needs to posit an eternal afterlife in order for agents to achieve perfect virtue. One might have thought that the task would be accomplished once one has had a few moments of respite from one's sensible self. We might wonder, in other words, what there is left for an agent to do once she is rid of inclination. But then we face a new problem from the other side of the highest good: once an agent is rid of sensibility and inclination, it is not clear how she can achieve happiness, at least as far as we understand it. Recall that Kant argues that we must postulate God as the author of laws of nature that guarantee the connection between virtue and happiness. But whatever happiness agents might enjoy once they've gotten rid of sensibility, it would not seem to be the sort of happiness brought about or conditioned by laws of nature.[13]

Finally, even if we ignore these problems, the argument seems to aim too high, doxastically speaking. Why assume that moral agency requires a guarantee of this object, such that practical belief in the existence of God and the immortality of the soul is warranted? Why, in other words, wouldn't something like practical belief in the mere possibility of the highest good and its postulates be sufficient?

These sorts of worries have led several commentators, including Paul Guyer, to argue for an interpretation according to which the highest good is best understood as a collective end, achievable over the course of human history. The highest good does not, in other words, describe a state of hoped-for existence for an individual, but rather the state of affairs brought about by the shared moral striving of the moral community. As such, Guyer sometimes calls this a 'unitary' account of the highest good, where happiness is "the condition that results from the successful pursuit of our ends, the promotion of happiness, not our personal happiness alone but the happiness of all insofar as that is compatible with the freedom of each."[14] Setting aside Kant's arguments regarding the postulates of pure practical reason in the Dialectic of the second *Critique*, there is ample textual evidence for such an

12 Guyer, Paul: "Kant, Mendelssohn, and Immortality." In Thomas Höwing (ed.), *The Highest Good in Kant's Philosophy*. Berlin/Boston 2016, 157–179, 171.

13 Cf. Guyer, Paul: "Kantian Communities." In *Virtues of Freedom: Selected Essays on Kant*. Oxford 2016, 276–302, 285.

14 Guyer, "Ends of Reason," 184.

interpretation of the highest good. In the *Critique of Pure Reason*, Kant likens the highest good to a 'moral world' in which morality is 'self-rewarding,' for example.[15]

Interpretatively, there is much to be said for this approach. First, it makes sense of the causal relationship that Kant suggests with regard to the relationship between virtue and happiness. For example, in the "Canon" of the first *Critique*, Kant acknowledges that "reason has causality with regard to freedom in general but not with regard to the whole of nature"[16] and identifies the highest good as "the world as it would be if it were in conformity with all moral laws (as it can be in accordance with the freedom of rational beings and should be in accordance with the necessary laws of morality) a moral world."[17] The 'world as it would be if it were in conformity will all moral laws' would be one in which agents observe perfect duty and perform imperfect duties, including imperfect duties to others. Since imperfect duties to others are duties to adopt the ends of others as one's own, and since accomplishing one's ends brings about — or perhaps just is — happiness, on Kant's understanding, there is a relatively straightforward connection between the two components of the highest good. Of course, this is not to say that any individual who is virtuous will also be happy, nor is it to suggest that moral action will be able to overcome natural misfortune. Still, the moral world described would be one in which happiness generally follows from virtue. Second, for reasons closely related to those above, this interpretation makes sense of Kant's claim that belief in at least one of the postulates of pure practical reason is a 'need of reason'.[18] If we have an obligation to do our part to bring about the highest good, then this would seem to place some demands on practical reason from an agential standpoint. Minimally, it would seem that some belief in the possibility of the highest good is required, along with whatever postulates necessary to secure this belief. While belief in the immortality of the soul is not required on this account, practical belief in laws of nature that link virtue and happiness is arguably warranted.[19] Finally, it is arguably a virtue of this interpretation that it makes the highest good a human project, and a project around which human institutions — both formal and less formal — can be organized. Not least of all, this includes the possibility that the organization of civil society can contribute to humanity's progress toward the highest good.[20]

15 Guyer, "Ends of Reason," 184, and "Kantian Communities," 283.

16 KrV, A 807/B 835.

17 KrV, A 808/B 836.

18 KpV, AA 05: 142–143.

19 For this reason, Guyer is reluctant to call even the interpretation of the highest good that places its fulfillment in this world a 'secular' notion of the highest good. See "Kantian Communities," 284.

20 Guyer: "Kantian Communities," 300.

As encouraging as all of this may be from the standpoint of a moral agent hoping that her moral strivings may eventually come to something and successfully produce happiness, it is likely to be cold comfort to us when we consider any individual virtuous life in isolation — including our own. To put the point in intentionally broad terms for now, there does seem to be something wrong or deficient about a virtuous life that is not also happy. Colloquially, at least, the point is made in terms of justice: it seems unjust or unfair that a virtuous life should not also be happy.

Of course, there a straightforward response open to the Kantian, namely to remind the agent concerned with 'justice' that the moral law doesn't aim at happiness. A major lesson of Kantian ethics, they might say, is that some degree — perhaps even a great degree — of personal sacrifice is demanded by the moral law. To insist that there is something unjust about a virtuous life that is not happy is just to allow sensibility to elbow its way in to our deliberations. It is essentially a psychological impulse, perhaps — to appropriate Bernard Williams' turn of phrase — we might even characterize it as a kind of 'squeamishness' about the demands of morality.[21]

There is an obvious sense in which all of this is true. But it is crucial to observe that much of what the Kantian might find objectionable about this demand for 'justice' is that there should be a causal link between virtue and happiness — either in the form of virtue naturally bringing about happiness (i.e., through laws of nature), or in the form of some divine reward for virtue. But if we set that assumption aside for a moment, I think there is still room for a set of important observations about the relation between virtue and happiness in the individual. In particular, since virtue and happiness are both ends of a rational yet sensible agent, there is a straightforward sense in which a life that is virtuous and happy is better than a life that is virtuous and unhappy. Seen from the standpoint of impartial reason the latter comes closer to a totality sought by practical reason. Thus, while the prospect of one's own happiness alone cannot serve as the object of moral action, the Kantian conception of practical reason will still judge there to be something lacking whenever a virtuous individual is not also happy to the degree that is consistent with that virtue. In what follows, I will attempt to unpack this point and try to discover what, if anything, follows from it.

21 See Williams, Bernard: "A Critique of Utilitarianism." In Smart, JJC: *Utilitarianism: For and Against*. Cambridge 1973, 77–151, 102.

2 Virtue, Happiness, and the Individual

Kant is not a Stoic: happiness matters to Kant, in the sense that it is a good in its own right — at least insofar as happiness is consistent with morality. Moreover, we should not, I think, presume that Kant's rejection of Stoicism is a reluctant concession to our sensible nature. To be sure, there are passages in which Kant seems to regret sensibility — for example, his claim in the *Groundwork* that the inclinations "are so far from having an absolute worth … that to be entirely free from them must … be the universal wish of every rational being",[22] or even the suggestion in the "Dialectic" discussed above, that our sensible nature makes virtue impossible at any moment of our existence. Guyer rightly observes that, seen in light of passages like these, the notion of the "composite" highest good within the individual "makes it sound as if virtue is our sole object as rational beings, but happiness our object of desire as finite beings, for which virtue must, somehow yet somewhat grudgingly, make room."[23] But I don't think that this is the view that Kant should take, nor do I think it is the view he always takes. In the *Religion*, for example, Kant claims that "Considered in themselves natural inclinations are good, i.e. not reprehensible".[24] And consider the following well-known passage from the "Dialectic" of the *Critique of Practical Reason*:

> [Virtue] is not yet, on that account, the whole and complete good as the object of the faculty of desire of rational finite beings; for this, happiness is also required, and that not merely in the partial eyes of a person who makes himself an end but even in the judgment of an impartial reason, which regards a person in the world generally as an end in itself. For, to need happiness, to be also worthy of it, and yet not to participate in it cannot be consistent with the perfect volition of a rational being that would at the same time have all power, even if we think of such a being only for the sake of the experiment.[25]

To my mind, this is one of Kant's better arguments concerning the 'composite' or individual highest good. He is wise to consider the matter third-personally, even, as he says, from "the standpoint of an impartial reason." This removes any danger that the concern for happiness is a heteronomous one. The point I take the passage to be making is fairly straightforward: given Kant's commitments in his practical philosophy — that we are end-setting creatures with the capacity to limit or constrain our pursuit of our particular ends according to the demands of pure practical

22 GMS, AA 04: 428.14–16.
23 Guyer: "Ends of Reason," 184.
24 RGV, AA 06: 58.01.
25 KpV, AA 05: 110.22–31.

reason — there are two very different things that constitute what is good for us. The first of these is virtue, and the second of these is happiness. Of course, the former is a necessary end, so it will always condition the latter. That is just to say that not all happiness is good; only happiness that is consistent with morality is good. But it is still good. From the standpoint of an impartial practical reason, a life that is maximally virtuous yet unhappy is thus less good — in the sense that it participates in less overall goodness — than a life that is maximally virtuous and also happy. Again, this is not to say that we ought always to seek own own happiness when we act, or even that we need some guarantee of eventual happiness in order to act morally. Both of those claims, I think, would be heteronomous. Nevertheless, we do recognize a kind of deficiency when a virtuous life is not also happy. Much of what makes this passage a good argument, on my view, is the modesty of its claims. In particular, there is no claim — yet — that happiness must follow from virtue, or even that agents should be rewarded for their virtue. Kant's arguments start to go off the rails once he begins to substitute these sorts of claims for the more modest observation that there is a deficiency in evidence when a virtuous person does not also "participate" in happiness. The best way, I think, to interpret the composite highest good, then, is not as a grudging concession to sensibility, but as something more like an embrace of it — insofar as it can be kept within the constraints of morality, of course.

So far, then, the point is merely that there is a coherent notion of the highest good for the individual, namely, the state of being maximally virtuous and simultaneously being as happy as such virtue allows. But it is also important to be clear about what this notion of the individual highest good does not include. First, there is no claim that this state of affairs exists or will exist. As we are all too well aware, it is a state that few — if any — achieve or enjoy, and Kant is under no delusions about this fact.[26] Second, it does not include a claim that there is much that any of us can do to bring about the highest good for ourselves or other individuals, aside from simply striving for virtue. In particular, there is no claim that it is our business to attempt to distribute happiness according to individuals' virtue, since we cannot know other people's motives in order to discern their virtue. And, in any case, Kant thinks that if doling out happiness according to virtue is anyone's job, it must be

26 Kant's most memorable acknowledgment of this fact is perhaps the discussion of the Spinozist in the *Critique of the Power of Judgment*: "Deceit, violence, and envy will always surround him, even though he is himself honest, peaceable, and benevolent; and the righteous ones besides himself that he will encounter will, in spite of all their worthiness to be happy, nevertheless be subject by nature, which pays no attention to that, to all the evils of poverty, illnesses, and untimely death, just like all the other animals on earth, and will always remain thus until one wide grave engulfs them all together…" (KU, AA 05: 452.20–27).

God's.[27] Third, despite what Kant will sometimes say, there is no philosophical reason to presume that there exists a causal relationship between individual virtue and happiness. If a person were to achieve the individual, composite highest good in the sense described so far, it would be to some degree because other agents are acting virtuously, but, presumably, it would also be largely a matter of luck. Nor is there any reason to think that virtue will cause or bring about happiness indirectly, via divine justice. Again, this is something that Kant will sometimes suggest, but it is not part of this basic notion of the individual highest good, according to which the best state of existence for creatures like us is one in which maximal happiness is conditioned by maximal virtue.

These observations bring us to a crucial distinction that, I think, Kant only sometimes takes heed of — much to the detriment of his arguments in the Dialectic, in particular. This is a distinction between object of practical reason and object of moral action. The object of practical reason refers merely to the totality sought by practical reason that we have been discussing thus far, and it is what Kant refers to when he says, for example, that "pure practical reason … seeks the unconditioned for the practically conditioned."[28] The object of moral action is what we can or should expect to result from virtuous action — whether this is understood in terms of a 'unitary' highest good, accomplished within the moral community, or whether it is understood in terms of divine reward for virtue.

Kant sometimes argues about the highest good as if the object of practical reason and the object of moral action were one and the same, but they are not. Kant's mistake, in other words, is to move too quickly from the wholly correct thought that virtue necessarily stands in a conditioning relationship with happiness to the further claim that virtue necessarily stands in a causal relationship with happiness. Of course, there is a perfectly legitimate reason to make this argumentative move when it comes to the virtue and happiness of the moral community, since agents acting virtuously will be treating others as ends in themselves and promoting the ends of others via the imperfect duty of beneficence. This will have a general causal tendency to promote the happiness of the moral community, taken as a whole.[29] But there is no similar or corresponding reason to assume a causal connection between virtue and happiness when it comes to the individual highest good. There, the relation is one of conditioning, not causation. Nevertheless, Kant clearly thinks that the relationship is causal in the "Dialectic":

27 *E.g.* V-Mo/Collins, AA 27: 450.30–35.

28 KpV, AA 05: 108.07–08.

29 Cf. "Ends of Reason," 184–185.

> Two determinations necessarily combined in one concept must be connected as ground and consequent, and so connected that this unity is considered either as analytic (logical connection) or as synthetic (real connection), the former in accordance with the law of identity, the latter in accordance with the law of causality. The connection of virtue with happiness can therefore be understood in one of two ways: either the endeavor to be virtuous and the rational pursuit of happiness are not two different actions but quite identical, in which case no maxim need be made the ground of the former other than that which serves for the latter; or else that connection is found in virtue's producing happiness as something different from the consciousness of virtue, as a cause produces an effect.[30]

In this passage, Kant seems only to consider two possibilities: either virtue and happiness are related to one another through a relation of identity, in which case their unity is analytic; or they have have a 'real' connection, in which case their unity is synthetic. As we have already seen, Kant rejects the ancients' way of proceeding, i.e. to assume a relation of identity, and insists instead on a 'real,' synthetic relation. But now Kant also insists that any synthetic connection between virtue and happiness must be causal. What is Kant's argument for this causal claim? Certainly, we can imagine two properties being related to one another without insisting on a causal relation between them: a mug might be blue and ceramic, but one property doesn't bring about the other. Presumably, Kant's claim in the passage above relies on the fact that we are not here considering two theoretical entities or properties, but rather two ends of practical reason. And since practical reason concerns itself with action — or causally bringing about certain effects in the world — any practical relation must be causal. As Kant explains,

> Now, this combination is (like every other) either analytic or synthetic. Since, as has already been shown, the given combination cannot be analytic, it must be thought synthetically and, indeed, as the connection of cause and effect, because it concerns a practical good, that is, one that is possible through action.[31]

The claim makes a certain amount of sense. When we think of the way that two actions are related to one another, we often think of the relation of means and ends. If my practical end is mastering the backstroke, then various other actions will be 'related' to this as means — for example, getting in the pool, and moving my legs and arms in a particular way. On the face of things, then, there is a case to be made for suggesting that non-identical actions relate to one another in a causal way. But we might add that my ends and actions can be related in other ways, too. If I want to master the backstroke, but I also want to meet a deadline for a separate project,

30 KpV, AA 05: 111.06–18, my emphases.

31 KpV, AA 05: 113.19–22, my emphasis.

then the things I need to do in order to master the backstroke may have to be put on hold until I finish the project. My end of finishing the project might, in other words, limit or condition my end of mastering the backstroke, at least temporarily. Or, to bring the point closer to the case at hand: I may need to put my end of mastering the backstroke on hold in order to fulfill a moral obligation — for example to rescue someone in distress.

Kant's insistence that virtue and happiness must be related to one another causally is, of course, a key premise in his argument for the postulate of the existence of God. Because virtue and happiness must relate to one another in this way, Kant argues that we must posit God as the author of laws of nature that secure this causal connection. But, as I have been suggesting, there is no need to insist on the causal premise. Notably, Kant does not always seem obviously to insist upon it himself. In his discussion of the highest good in the Canon of the Critique of Pure Reason, Kant sketches the relation as follows:

> Leibniz called the world, insofar as in it one attends only to rational beings and their interconnection in accordance with moral laws under the rule of the highest good, the realm of grace, and distinguished it from the realm of nature, where, to be sure, rational beings stand under moral laws but cannot expect any success for their conduct except in accordance with the course of nature in our sensible world. Thus to regard ourselves as in the realm of grace, where every happiness awaits us as long as we do not ourselves limit our share of it through the unworthiness to be happy, is a practically necessary idea of reason.[32]

Kant's description of the 'realm of grace' in this passage does not seem to suggest a strong causal connection between virtue and happiness. Rather, he describes it as a realm 'where every happiness awaits us' [wo alle Glückseligkeit auf uns wartet] as long as we do not make ourselves unworthy of happiness. To be sure, the notion of happiness 'awaiting' us is potentially ambiguous: it could 'await us' qua reward for virtue, in which case Kant would be suggesting a kind of causal relation between virtue and happiness. More plausible, to my mind, is an interpretation according to which the 'realm of grace' is simply a maximally happy place, in which agents are free to enjoy happiness, so long as they are worthy of it. On that interpretation, there is no causal connection between virtue and happiness. The source, as it were, of happiness is not related to a person's virtue, either naturally or through divine reward. Nevertheless, a person's virtue will condition that happiness, i.e. make her deserving of it.

32 KrV, A 812/B 840.

The passage from the *Critique of Pure Reason* is not an outlier. Kant reportedly says something very similar in the second set of *Mrongovius lectures*, given around the time of the publication of the *Groundwork*:

> The mistake [of the ancients] in this was that they tried to squeeze the two parts into one, when in fact they are entirely different; for the one indicates the worth of the person, whereas the other indicates the worth of the condition. The two are heterogeneous and both have therefore their own particular sources. But there must be a link; otherwise they can never be together. Now, this link is religion. The ancients tried to do without religion, but it did not work. Only through religion may I hope that someone who makes himself worthy of happiness can actually be happy.[33]

In this passage, as in the previous passage, there does not appear to be any mention of a causal connection between virtue and happiness. Of course, since the source of the passage is a set of transcribed lecture notes, it is important not to put too much weight on particular words that Kant did or did not say. Still, the general sense of the passage is interesting for its similarity with the discussion of the ancients in the Dialectic, and for the way that it appears to diverge from that discussion. As in the "Dialectic," Kant observes that the ancients made the mistake of identifying virtue with happiness — trying to squeeze two parts into one. Notably, however, Kant also observes that virtue and happiness have two different sources, a claim that is, to be sure, consistent with a causal claim, but perhaps friendlier to a more moderate claim that virtue and happiness are two very different ends, each with its own source. Despite the fact that virtue and happiness have two different sources, Kant insists that there 'must be a link' between them, offering what sounds like a tautology for an argument: if there is no link, then they can never be together. Again, we don't know if this is what Kant actually said, but a charitable way of interpreting Kant's claim would be to point to the totality that practical reason seeks, a totality in which virtue and happiness are found together. Since virtue and happiness have distinct sources — at least as far as the individual is concerned — Kant thinks we need to look to religion for some hope of an association between the two. Why or whether we need such hope is its own question, which I consider in the next section.

When it comes to the hoped-for relation between virtue and happiness, I suspect that the argumentative and interpretative waters are muddied somewhat by the fact that Kant does consistently observe that there is no reason to think that there exists a causal connection between virtue and happiness for the individual agent in this world. The passage from §87 of the *Critique of the Power of Judgment*

33 V-Mo/Mron II, AA 29: 600 (translated Jens Timmermann and Stefano Bacin).

quoted in the footnote above is certainly an example, as is the following observation in the *Critique of Pure Reason*:

> But since the obligation from the moral law remains valid for each particular use of freedom even if others do not conduct themselves in accord with this law, how their consequences will be related to happiness is determined neither by the nature of things in the world, nor by the causality of actions themselves and their relation to morality; and the necessary connection of the hope of being happy with the unremitting effort to make oneself worthy of happiness that has been adduced cannot be cognized through reason if it is grounded merely in nature, but may be hoped for only if it is at the same time grounded on a highest reason, which commands in accordance with moral laws, as at the same time the cause of nature.[34]

Still, to observe correctly that there is no causal connection to be found between two ends with distinct sources is not to say that these must be somehow connected causally. There is a difference between two sorts of arguments, and I suspect they often bleed into one another in Kant's writing and our interpretation of him. The first kind of argument begins with the premise that individual virtue and happiness must be causally connected and then observes that such a causal connection does not appear to hold in the order of nature. The second kind of argument merely observes that virtue and happiness are both goods — virtue, of course, being necessarily good, and happiness only conditionally good — and observes, with some regret, that there is no predictable, causal, way to bring the latter about by pursuing the necessary end of morality. Indeed, pursuing the end of virtue will often have just the opposite result when it comes to the pursuit of happiness. The first type of argument makes the accomplishment of happiness a project for an afterlife; the second makes the accomplishment of happiness a matter of luck.

There are, in other words, two distinct possible interpretations of Kant's claim in the *Religion* that "it cannot be a matter of indifference to reason how to answer the question, What is then the result of this right conduct of ours?"[35] On the first interpretation, the question whose answer cannot possibly be a matter of indifference is about the causal outcome of our moral actions. Of course, there is an obvious Kantian sense in which agents should be indifferent to the answer to this question, so long as it is about the happiness we can expect to enjoy individually as a result of our moral actions. The causal interpretation of the question makes sense, in other words, when we understand it to be a question about the happiness we can expect to produce in the moral community as a result of our moral actions. But there is another way of understanding the question whose answer 'cannot be a

34 KrV, A 810/B 838.

35 RGV, AA 06: 05.03–04.

matter of indifference', and this is a question about whether we can, as individuals, ever reasonably hope to be happy and virtuous simultaneously — whatever the source of that happiness. This, in other words, is a question about how close to the complete good of practical reason any individual can hope to come.

Of course, it is less than obvious what follows from the preceding observations about the object of practical reason when it comes to what we can reasonably conclude or hope for when it comes to metaphysical questions concerning the existence of God or the immortality of the soul. Because the relationship between personal virtue and happiness is conditioning, and not causal, there does not seem to be any separate duty to bring about the personal highest good, and so the conclusions involving the "need of reason" mentioned earlier do not follow. Still, there may be a need of reason specific to the personal highest good, understood as a totality of practical reason where virtue conditions happiness. We turn to that question next.

3 The Need of Reason

Closely related to the preceding questions about how, precisely, to understand Kant's notion of the highest good is another set of questions about what Kant's account of the highest good gives us license to believe, or at least hope, on practical grounds. As Kant explains in the *Critique of Practical Reason*,

> [A] need of pure practical reason is based on a duty, that of making something (the highest good) the object of my will so as to promote it with all my powers; and thus I must suppose its possibility and so too the conditions for this, namely God, freedom, and immortality, because I cannot prove these by my speculative reason, although I can also not refute them.[36]

Speculative reason, as Kant takes himself to have shown decisively in the *Critique of Pure Reason*, cannot make claims about the existence — or non-existence — of God, freedom, and the immortality of the soul. However, Kant now argues, his arguments regarding morality and the highest good give us practical grounds for belief in them. Indeed, Kant has already argued in the Preface to the *Critique of Practical Reason* that we have practical knowledge of freedom, since this follows immediately from our awareness of morality's bindingness.[37] Now Kant argues that practical belief in God and the immortality of the soul can be supported by the needs of pure practical reason.

36 KpV, 05: 142.19–24.

37 KpV, AA 05: 04.07–10.

This 'need of reason' argument, as Kant presents it, clearly relies on the assumption that virtue and happiness are related to one another causally. The version of the highest good described in this passage is both the object of my will and something that ought to be promoted with all of my powers. As Kant explains in his response to Thomas Wizenmann, it is precisely because we have a duty to bring about the highest good that this object of our will is not a mere wish, but rather a need of reason.[38] Duty tells us that we have an obligation to promote the highest good, so it must be possible. Hence, we are licensed to have a practical belief in those things that make it possible — in particular, the immortality of the soul and the existence of God. There are many questions that we can raise about the argument — for example about the various modalities of belief that Kant mentions: Is belief in God as the author of laws of nature required to support the possibility of the highest good, or might something like belief in the mere possibility of such an author be sufficient? However we come down on questions like these, it is still the case that the argument only works if our understanding of the highest good is that of an object of moral action, not just an object of practical reason. To put a finer point on things, if we understand the highest good in terms of a totality that practical reason necessarily aims at, it seems like it will be much closer to the kind of hope or wish that Wizenmann accuses Kant of placing at the foundation of his argument. Practical reason must, necessarily, aim at virtue, and it does, definitionally, aim at happiness. But if we think of these as causally disconnected, then any duty to bring about the highest good cannot go beyond the duty that we already knew we had, namely, to be as virtuous as possible.

This is why I tend to think that the need of reason passage works better — but not perfectly — with the 'unitary' or immanent conception of the highest good. On that account, we can understand it as our duty to bring about both components of the highest good, i.e. we can understand the highest good as the object of moral action. Of course, the argumentative payoff of this interpretation is less impressive, since there is no need to posit the immortality of the soul if progress can happen over the course of many generations. But my sense is that it is enough to support a practical belief that there is a general causal connection between virtuous action and happiness. On one level, such a causal connection is not very mysterious: we aim at bringing about happiness when we adopt the ends of others as our own, for example. Still, I think, there is room for a kind of faith that our actions will be effective, i.e. that the causal relation between adopting and acting for the sake of another's ends and their happiness will generally exist.

38 KpV, AA 05: 143n.

However, as I have been arguing, Kant is not really at liberty to say that the connection between virtue and happiness in the individual is a causal connection. He is right to say that it is an object of practical reason in a certain sense. Specifically, it describes the complete good for an individual, and thus constitutes a coherent object for practical reason, which is bound to seek this kind of totality. But there isn't really any sense in which we should think that virtue causes happiness in the individual, so any metaphysical conclusions about what must be the case in order to make possible that causal connection are blocked. Now, we might think that this leaves us with a rather underwhelming conclusion regarding the individual, 'composite' highest good. Yes, perhaps there is a sense in which it is a coherent goal for practical reason to strive toward, but if this is really all there is to it, there isn't much we can conclude from this observation. Nevertheless, I think there are Kantian reasons to be suspicious of a universe that is wholly indifferent to the highest good, understood individually. Perhaps a somewhat lighthearted example may help illustrate this thought.

Imagine two teams engaged in a competitive game. Two aims will guide the conduct of each team. First, each team must follow the rules of the game. But there is a rather obvious sense in which the teams are not there simply to follow the rules of the game; they are also trying to win the game. These two aims model the aims of practical reason. The first — following the rules of the game — is necessary and conditions the other. A team that scores more points by cheating hasn't won, or at least this is how an 'impartial reason' will judge the matter. But, again, the teams are each there to win, so we can say that the complete good for each team is to play according to the rules of the game and win the game. Though the necessary end of playing by the rules conditions the end of winning the game, it does not bring it about that a team will win the game. Nor should we expect it to, since there is no necessary causal link between playing by the rules and winning the game: these are separate aims. Nor is it the case that either team needs a guarantee of winning in order to play the game. Indeed, in a game that doesn't allow a tie, each team knows that one team will not win. Still, I think there is a fairly fundamental sense in which it would not make sense for a team to play the game at all if there were reason to suspect that they could never win. Suppose a referee has been bribed always to rule against their team, for example. 'What's the point of playing,' they might say, 'if we don't stand a chance of winning?' Of course, one hears such things when teams are in a slump or down on their luck, but that is probably hyperbole: in those cases, there is still a chance of winning, even if it is distant and would

require a good deal of effort or luck.[39] The case I have in mind is a case in which a team has knowledge — sufficient objective grounds for assent — that they cannot win, despite their best efforts, and despite their playing by the rules.[40] I think it would be close to impossible, if not simply nonsensical, to play the game under these conditions.

I suggest the 'need of reason' as it relates to the individual highest good is like this. Just as there is no reason to think that following the rules of a game will bring about a team's victory, there is no reason to expect that happiness will follow from virtue — either via natural causal laws, or through divine reward. Still, practical reason tells us that there is a complete good for creatures like us, and we necessarily act in a way that seeks this complete good. Thus, from the standpoint of a practical reason that seeks this totality, it would be a serious problem if we knew that the complete good were impossible. It would be impossible, if not nonsensical, to act if we knew that the world were wholly inhospitable to the pursuit of the individual highest good, understood as this totality. To be clear, this is miles away from the claim that we could only act if our own happiness were assured. But it would be more than just discouraging if we lived in a world whose laws were set up in such a way that virtue could not be coupled with happiness, even if only fortuitously. This, at a minimum, is what I take Kant to be saying in the *Mrongovius lectures* when he says there must be a link between virtue and happiness.

In those lectures, Kant says that it is the role of religion to ensure this link between virtue and happiness. In light of this discussion of the need of reason, it is not wholly obvious what this might mean. If it means that we have to hope that virtuous individuals will achieve happiness in an afterlife, then this may not be altogether satisfying, since, as we have already noted, it is not clear how happiness can be part of a non-sensible afterlife. Perhaps Kant means that it is the role of religion to give us some rational belief that the world is at least hospitable to the virtuous agent's happiness, i.e. that when it comes to the laws of nature, things are set up so that virtue and happiness can coexist. Again, if we think of the highest good in the individual as a kind of totality as opposed to a causal relationship between virtue and happiness, then the stakes are somewhat lower. We do not have to have rational belief in laws of nature that ensure a causal connection between virtue and happiness in the individual. We just need a belief in laws of nature that don't get in the way of these two things coexisting. This is not a minor point: it is precisely

39 Alternately, in such cases, we might make sense of the team forging ahead despite the impossibility of winning if we say that they are playing for the sake of their teammates, or the love of the game, *etc.*

40 Chignell, Andrew: "Belief in Kant." In *Philosophical Review* 116 (3), 2007, 323–360.

because Kant's moral system decouples virtue and happiness when it comes to the ground of moral obligation that there is some need to seek this reassurance from the perspective of practical reason, considered as a whole.

To go beyond what Kant says on the matter, I think it is possible to extend the observation to human institutions. As we have already noted, it is clearly not our business to reward individuals for their virtue, or to try to make sure that virtuous people are happy and vicious people are unhappy. Nevertheless, seen from the standpoint of the needs of reason associated with the individual highest good, it would be practically undermining to devise institutions that make the world inhospitable to the happiness of virtuous agents — think back to the corrupt referee, for example.

4 Conclusion

It bears emphasizing that I do not take myself to be doing strict exegesis in this discussion: that is, I do not think that the arguments I've presented are able to 'make sense' of everything that Kant says about the highest good, and untangle and confusions or apparent problems with his arguments. For example, Kant clearly thinks that there is a causal connection between individual virtue and individual happiness in the Dialectic of the *Critique of Practical Reason*, and I think there is no reason for him to help himself to this premise. There is also a clear sense in which the account of the individual highest good I have been offering here works less well alongside Kant's claims that we have a duty to pursue the highest good,[41] or — relatedly — the claim that the object of the highest good can be understood already as part of the ground of moral action.[42] It is clear to me that the systematic, this-worldly interpretation of the highest good does a much better job of making sense of these claims. My main goal in this paper is simply to argue that we should not allow the obvious strengths of the systematic account to let us discard an individualist account altogether. There are good Kantian reasons for maintaining a pared-down, individualist notion of the highest good — understood as a totality of practical reason that every rational and sensible agent wants to participate in. At a minimum, this account of the highest good can give us a sense of what a good life looks like for a Kantian agent. Beyond this, it may provide Kantian grounds for belief in laws of nature that are at least not hostile to the fortuitous coupling of virtue and happiness, and it may give us good reason not to devise institutions that are hostile to happiness.

41 KpV, AA 05: 126.01.

42 KpV, AA 05: 109–110.

References

Beck, Lewis White (1960). *A Commentary on Kant's Critique of Practical Reason*, Chicago.
Chignell, Andrew (2007). "Belief in Kant." *Philosophical Review* 116 (3), 323–360.
Guyer, Paul (2005). "Ends of Reason and Ends of Nature." In *Kant's System of Nature and Freedom*, Oxford, 169–197.
Guyer, Paul (2007). "Naturalistic and Transcendental Moments in Kant's Moral Philosophy." *Inquiry* 50, 444–464.
Guyer, Paul (2016a). "Kant, Mendelssohn, and Immortality." In *The Highest Good in Kant's Philosophy*, ed. by Thomas Höwing, Berlin/Boston, 157–179.
Guyer, Paul (2016b). "Kantian Communities." In *Virtues of Freedom: Selected Essays on Kant*, Oxford, 276–302.
Timmermann, Jens (2022). *Kant's Will at the Crossroads*, Cambridge.
Williams, Bernard (1973). "A Critique of Utilitarianism." In *Utilitarianism: For and Against*, ed. by J.J.C. Smart, Cambridge, 77–151.