Kant's "Moral Proof": Defense and Implications

Michael Baur

Kant's "moral proof" for the existence of God has been the subject of much criticism, even among his most sympathetic commentators.[1] According to the critics, the primary problem is that the notion of the "highest good," on which the moral proof depends, introduces an element of contingency and heteronomy into Kant's otherwise strict, autonomy-based moral thinking. In this paper, I shall argue that Kant's moral proof is not only more defensible than commentators have typically acknowledged, but also has some very interesting implications as well (e.g., the moral proof is "circular" and thus implicitly self-validating). My account shall proceed in five stages:

1. Preliminary Discussion of the Moral Proof
2. The Argument of the Moral Proof
3. Criticism of the Moral Proof
4. Defense of the Moral Proof
5. Implications of the Moral Proof: Circularity and Self-referentiality

1. Preliminary Discussion of the Moral Proof

Kant is rightly famous for his thoroughgoing critique of speculative, or theoretical, proofs for God's existence, including the ontological, cosmological, and teleological (or what Kant calls the physico-theological) proof.[2] My aim here is to focus not on Kant's critique of speculative proofs, but rather on his argument for the so-called "moral proof" of God's existence. I shall begin with three preliminary remarks.

First, one should not think that Kant's moral proof is intended as a second-best alternative to a speculative or theoretical proof that would have been preferable to have, if one were only available. For Kant, theoretical reason is subordinate to practical reason, and the interests of practical reason itself require that there be no theoretically valid proof for God's existence. For if it were possible definitively to prove God's existence on the basis of theory alone, the autonomy of our practical, moral activity would be compromised. As Kant writes, if we had speculative knowledge of God's existence, then:

> *God and eternity with their awful majesty* would stand unceasingly *before our eyes* (for what we can prove perfectly holds as much certainty for us as what we are assured of by our sight). Transgression of the [moral] law would, no doubt, be avoided: what is commanded would be done; but . . . the spur to activity in this case would be promptly at hand and *external* . . . . [M]ost actions conforming to the law would be done from fear, only a few from hope, and none at all from duty, and the moral worth of actions, on which alone in the eyes of supreme wisdom the worth of the person and even that of the world depends, would not exist at all.[3]

Thus Kant's famous statement—that ". . . I had to deny knowledge in order to make room for faith" (*CPR*, B xxx)—is not one of regret, but of emancipation: the failure of theoretical proofs for God's existence opens a space for genuinely moral—and that means autonomous—practical activity.

Secondly, Kant's "moral proof" is actually misnamed. The proof does not aim to demonstrate that we can *know* of God's actual, independent existence, but only that we are rationally justified—on moral, if not theoretical, grounds—for *believing* that God exists. It does this by showing that—given the requirements of pure practical reason—we must believe in God; without the practical postulate of God's existence, we would have to think of our moral activity, and thus the moral law upon which it is based, as empty and pointless.

Thirdly, while Kant's proof is grounded on moral—and not theoretical—principles, it is not simply an argument by means of wishful thinking. One of Kant's own contemporaries, Thomas Wizenmann, had compared the moral proof to the wishful thinking of "a *man in love*, who, having fooled himself into an idea of beauty that is merely a chimera of his own brain, would like to conclude that such an object really exists somewhere."[4] Kant responded by acknowledging that this criticism would have been well-founded if the practical necessity of the postulate of God were related to a person's *subjective* wishes or inclinations; but the criticism ultimately fails since the practical postulate is grounded only on the *objective* (necessary and universal) requirements of practical reason itself:

> I grant that he is perfectly correct in this, in all cases where the need is based upon *inclination*, which cannot necessarily postulate the existence of its object even for the one affected by it . . . . But in the present case it is a *need of reason* arising from an *objective* determining ground of the will, namely the moral law, which necessarily binds every rational being and therefore justifies him apriori in presupposing in nature the conditions befitting it and makes the latter inseparable from the complete practical use of reason (*CPrR*, 5:144).

2. The Argument of the Moral Proof

Kant's moral proof is discussed in each of his three critiques, but it receives its clearest and most complete formulation in the second critique, or *Critique of Practical Reason*. My commentary will thus focus on the second critique, with passing reference to Kant's other works as well.

The moral proof begins with the notion of the "highest good," and a distinction between two possible senses of the term: the highest good as the "supreme" good and the highest good as the "most complete" or "most perfect" good. A morally good or virtuous will (one that acts out of pure duty or respect for the moral law) is supremely and unconditionally good; however, it does not follow that a good or virtuous will is the only possible good. While virtue is the "supreme" good in the sense that it is the unconditioned condition of all other goods, it does not follow that it is the "most complete" or "most perfect" good. What is required for the "highest good" in the sense of "completeness" and "perfection" is not just a good or virtuous will, but also happiness in proportion to such virtue. A world in which there was virtue, but no happiness in proportion to it, would not be a perfectly or completely good world.[5]

Before proceeding any further, it is important to address here what seems to be an obvious criticism, one first raised by Kant's contemporary, Christian Garve.[6] The criticism has to do with the fact that Kant's notion of the highest good includes happiness as one of its elements. The problem is that Kant's own autonomy-based moral thinking seems to imply that happiness should be entirely irrelevant to our moral concerns. Kant responds to this criticism in his essay on "Theory and Practice," arguing that the emphasis on autonomy need not be understood to imply that happiness is not relevant or is not genuinely good. Of course, the unconditional character of the moral law does entail that one's desire for happiness should never stand in the way of one's doing what is morally right. Nevertheless, this in itself does not entail that moral goodness is the only possible goodness or that happiness is not itself good. It entails only that happiness is a *conditional* good; but when and if the conditions of this conditional good are satisfied, happiness can be a genuine good. Thus Kant's notion of the "highest good," in itself, poses no problems for the rest of his moral theory. The notion of the "highest good" involves a combination of two heterogeneous elements, virtue and happiness, each of which can be genuinely good; the crucial point, however, is that one good (a morally good or virtuous will) is unconditionally good, while the other (happiness) is only conditionally so.[7]

In a similar vein, one can defend Kant here by reference to an important terminological distinction that Kant employs. For Kant, the *determining ground* of our moral activity is simply that which serves as the incentive or

reason for one's undertaking a morally significant action. The *object* of our moral activity is defined as the effect or result to be brought about through the freedom of our will (*CPrR*, 5:57). Now for Kant, the only proper *determining ground* of our moral activity is respect for the moral law (or the virtuous will itself, regarded as legislating universal law). But the proper *object* of a morally good will need not be restricted to mere respect for the moral law or to the mere actualization of a virtuous (but otherwise ineffectual) will. For Kant, then, the highest good—as the combination of virtue and happiness proportionate to it—can be the full and complete *object* of pure practical reason, yet without being its *determining ground* (*CPrR*, 5:109). We shall only later be in a position to examine whether Kant's response to Garve and his distinction between the will's *determining ground* and its *object* can ultimately protect Kant against other possible criticisms.

Now in addition to saying that the highest good can be a genuinely good object of the will, Kant goes on to claim that we do, in fact, have an obligation or duty to promote it. A world in which a person is "in need of happiness and also worthy of it," but still does "not partake of it" is a morally defective world, one that "could not be in accordance with the complete volition of an omnipotent rational being" (*CPrR*, 5:110). Now for Kant, to have a good or virtuous will is the same as to be worthy or deserving of happiness (*CPrR*, 5:110); thus our obligation to promote the highest good—which involves a proportionality between virtue and happiness—is simultaneously an obligation to promote a proportionality between desert and reward. But a proportionality between desert and reward is the same as justice (*CPrR*, 5:115, and 5:123). Therefore, our obligation to promote the highest good is equally an obligation to promote justice. Furthermore, since virtue is an effect of our freedom alone, and happiness is an effect of natural causes insofar as they relate to our desires and inclinations, it follows for Kant that the obligation to promote the highest good is also an obligation to bring about a harmony between freedom and nature. Thus Kant explicitly speaks of the highest good as "the Kingdom of God in which nature and morality come into harmony with one another" (*CPrR*, 5:128).[8]

Now our obligation to promote the highest good raises a difficulty, and it is this very difficulty that leads Kant to conclude that belief in God is rationally justified and necessary. That is, Kant's explanation and overcoming of the difficulty associated with the highest good constitutes the substance of his moral proof, to which we now turn.

For Kant, we have an obligation to promote the highest good, understood as a combination or harmony of two heterogeneous elements: virtue and happiness. But because of the heterogeneity of these two elements, any connection between virtue and happiness is not merely analytic but synthetic.[9] Furthermore, this synthetic connection can be conceived in only two

possible ways. Either the desire for happiness is the ground of virtue, or else the maxim of virtue is the ground of happiness (*CPrR*, 5:113). Now the first option is utterly impossible, for that would destroy the purity and autonomy of practical reason by locating the determining ground of the will in the desire for happiness. The second option, however, is also impossible: the capacity to experience happiness is not directly dependent on the moral intentions of one's will, but rather on causes and effects in nature (including one's own desires and inclinations, etc.). Thus, while we have an obligation to promote the highest good, there seems to be no ground that could possibly guarantee the requisite connection between virtue and happiness. Morally good people suffer, while morally bad people thrive, and our moral goodness alone seems utterly powerless to bring about the requisite harmony between virtue and happiness.

Now if we are morally obligated to bring about such a harmony, but cannot do so through the goodness of our will alone, then any moral imperative that obliges us to do so must be null and void. For an obligation that obliges us to achieve what is beyond our control cannot be an obligation at all.[10] In turn, the emptiness of this obligation undermines the moral law itself, insofar as this obligation and the moral law are connected to one another. As Kant writes:

> . . . no necessary connection of happiness with virtue in the world, adequate to the highest good, can be expected from the most meticulous observance of moral laws. Now, since the promotion of the highest good, which contains this connection in its concept, is an apriori necessary object of our will and inseparably bound up with the moral law, the impossibility of the first must also prove the falsity of the second. If, therefore, the highest good is impossible in accordance with practical rules, then the moral law, which commands us to promote it, must be fantastic and directed to empty imaginary ends and must therefore in itself be false (*CPrR*, 5:113-114).

Because our obligation to promote the highest good is connected to the moral law and because an obligation that requires what is beyond our control can be no obligation at all, Kant argues that there must be some way in which we can think it possible to promote the highest good through our own moral agency—otherwise, the moral law would lose its authority and binding force.

Kant begins to address the problem by pointing out that the initial absence of any guaranteed connection between virtue and happiness leads to an insuperable difficulty *only if* one first assumes that the ground of any such connection must reside in the moral activity of *finite* wills. The moral activity of such wills, as finite, necessarily presupposes the pre-existence of a given natural world (including external objects and each person's natural inclinations) *upon which* such activity is exercised. In other words, the

finitude of such moral agents entails that the whole natural world upon which their moral activity is exercised is itself not *already* a product of their *own* moral activity (*CPrR*, 5:124). But since the natural world upon which such moral activity is exercised is itself not a product of this very moral activity, and since there seems to be no other source from which the natural world might acquire moral significance or direction, there seems to be no conceivable ground that can guarantee the complete harmony between virtue and happiness, freedom and nature. In other words, nature is at first simply "given" as indifferent and unrelated to the moral activity of finite rational agents. And because nature, so considered, is morally indifferent, there can be no guarantee that our finite moral activity can ultimately bring about the highest good as a harmony between virtue and happiness, freedom and nature. Accordingly, any obligation to promote the highest good seems to require something that is beyond our control, and so the obligation—along with the moral law connected to it—appears to be null and void.

As Kant goes on to argue, this conclusion can be avoided only if one assumes the existence of a will that is not finite like our own, and thus not dependent on a pre-existing natural world—that is, only if one assumes the existence of a good and all-powerful God who created the natural world, and indeed created it such that it is not wholly indifferent to our moral purposes but completely conformable to them insofar as they are morally virtuous.[11] Thus even though there is no necessary connection between *my* finite moral activity and the causes and effects that occur in the natural world, I can *think* of this connection indirectly, as mediated and guaranteed by the will of "an intelligible author of Nature" (*CPrR*, 5:115). As Kant summarizes the argument:

> . . . the acting rational being in the world is . . . not also the cause of the world and of nature itself. Consequently, there is not the least ground in the moral law for a necessary connection between the morality and the proportionate happiness of a being belonging to the world as part of it and hence dependent upon it, who for that reason cannot by his will be a cause of this nature and, as far as happiness is concerned, cannot by his own powers make it harmonize thoroughly with his practical principles. Nevertheless, in the practical task of pure reason, that is, in the necessary pursuit of the highest good, such a connection is postulated as necessary: we *ought* to strive to promote the highest good (which must therefore be possible). Accordingly, the existence of a cause of all nature, which contains the ground of this connection, namely of the exact correspondence of happiness with morality, is also *postulated* . . . . Therefore, the highest good in the world is possible only insofar as a supreme cause of nature having a causality in keeping with the moral disposition is assumed (*CPrR*, 5:124–125).

Thus our belief in the existence of God is not only justified but also required as a matter of practical reason, insofar as we have an obligation to promote the highest good:

> Now, it was a duty for us to promote the highest good; hence there is in us not merely the warrant but also the necessity, as need connected with duty, to presuppose the possibility of this highest good, which, since it is possible only under the condition of the existence of God, connects the presupposition of the existence of God inseparably with duty; that is, it is morally necessary to assume the existence of God (*CPrR*, 5:125).

While, for Kant, it is necessary to believe in God's existence, such necessity does not entail any form of heteronomy, since this necessity is grounded in the self-legislated requirements of practical reason itself. The postulate of God's existence concerns only "the possibility of the highest good, though not for the sake of some arbitrary speculative design but only for the sake of a practically necessary end of the pure rational will, which does not here choose but rather obeys an inexorable command of reason" (*CPrR*, 5:143). Thus our faith in God is fully consistent with Kant's autonomy-based thinking; for Kant, "pure reason alone . . . is the source from which [our rational faith] springs" (*CPrR*, 5:126).

While Kant's moral proof does not depend on any externally-derived, dogmatically-imposed content, the notion of God that it points to is not entirely contentless; the moral proof does not lead to an empty deism of practical reason. Instead, the proof entails that God, as the guarantor of a harmony between freedom and nature, must possess a number of specific perfections. As Kant writes, the will of God:

> must be omnipotent, so that all of nature and its relation to morality in the world are subject to it; omniscient, so that it cognizes the inmost dispositions and their moral worth; omnipresent, so that it is immediately ready for every need that is demanded by the highest good for the world; eternal, so that this agreement of nature and freedom is not lacking at any time, etc. (*CPR*, A815/B843).[12]

Finally, the moral proof—if it works—allows Kant to combine religious thinking with morality, but without accepting the heteronomy inherent in all divine command theories of morality:

> . . . through the concept of the highest good as the object and final end of pure practical reason, the moral law leads to religion. Religion is the recognition of all duties as divine commands, not as sanctions, i.e., arbitrary and contingent ordinances of a foreign will, but as essential laws of any free will as such (*CPrR*, 5:129).

Accordingly:

> . . . the Christian principle of morality is not theological and thus heteronomous, being rather the autonomy of pure practical reason itself, because it does not make the knowledge of God and His will the basis of these laws but makes such knowledge the basis only of succeeding to the highest good on condition of obedience to these laws (*CPrR*, 5:129).

In short, for Kant, what we take to be divine commands are not binding on us simply because they are divine commands; rather, we regard certain imperatives as divine commands because they are already binding on us in accordance with the self-legislated imperatives of our own autonomous practical reason.[13]

3. Criticism of the Moral Proof

While Kant's moral proof may be defensible in general, there is a seemingly devastating criticism that can be leveled against it; one that has confounded even the most sympathetic Kant-interpreters.[14] Let us assume—for the sake of argument—that one accepts Kant's response to Garve: there is a difference between the *determining ground* and the *object* of the will, and happiness can be a genuine good, even though it is not an unconditional good. Even so, it does not follow that we have a *moral*—as opposed to merely *prudential*—obligation to promote the highest good as constituted by both virtue and happiness. Any concern that we might have regarding happiness, including any concern over its possible connection to virtue, bears no necessary connection to the *moral* law, as Kant claims it does.[15] Thus our inability to bring about the highest good on our own, in the absence of God, does not entail our inability to bring about that which we are *morally* obliged to bring about, and thus does not entail that the *moral* law is null and void without the postulate of God. Therefore, it is not necessary—as Kant says it is—to postulate the existence of God in order to conceive the possibility of our fulfilling our *moral* obligations.

Another way to express the difficulty is to point to the fact that Kant's moral proof seems viciously circular: our apparent inability to bring about the highest good on our own, in the absence of God, can cut one of two ways, and it cuts in favor of Kant's moral proof only if one already assumes an answer to what is genuinely at issue. On one side, it is possible to argue—as Kant does—that we can never have a *moral* obligation to do that which is beyond our control; but the promotion of the highest good is something beyond our control, except on the supposition that God exists; therefore, as a matter of practical necessity, we must accept the postulate of God's existence. On the other side, it is possible to argue—beginning with Kant's very

own premise—that we can never have a *moral* obligation to do that which is beyond our control; but secondly, the promotion of the highest good is something beyond our control; therefore, we must not have a specifically *moral* obligation to promote the highest good (even though we may have a prudential or eudaemonistic obligation to do so).

Now in both of these conflicting arguments, it is assumed that we must think it possible to achieve that which we are *morally* obliged to do. The crucial issue, then, is whether we are morally obliged to promote only virtue (which we are capable of promoting through our moral willing alone) or the highest good (including happiness, which is beyond the power of our moral willing alone). Kant's moral argument succeeds only if one already assumes that we have a *moral* duty to promote that which is seemingly beyond the power of our moral willing alone. But it is this very assumption that can be questioned on Kant's own moral grounds. Why must one assume that we have a *moral*—and not merely *prudential*—duty to promote the highest good as something that is beyond our direct control?

Unfortunately, Kant merely states—without explanation—that a concern over happiness and its proportionate relation to virtue is necessarily related to the moral law (*CPrR*, 5:113-114). What makes matters worse, it seems that Kant's claim that we have a *moral* duty to promote the highest good (which includes "happiness" proportionate to virtue) threatens to make our moral activity heteronomous; after all, the desirability of happiness seems related more to our passions and inclinations, and not to the *moral* law as such.

One possible response to this problem might be based on Kant's explanation that the notion of happiness contained in the "highest good" is not merely one's own *subjective* happiness, but rather happiness from the perspective of rationality as such. This happiness is not merely happiness "in the partial eyes of a person who makes himself his end but even in the judgement of impartial reason, which impartially regards persons as ends in themselves" (*CPrR*, 5:110). Unfortunately, this strategy alone does not respond to the fundamental problem. For the objection is not just that *my* happiness, but rather that happiness at all—whether from the perspective of one person or reason as such—should be connected to our *moral* obligations. But why—on Kantian grounds—should we have a specifically *moral* obligation to be concerned about happiness at all, when possible happiness in the world is a result or consequence that is beyond our control as *moral* agents?

Later in this paper, I shall seek to show that Kant's position is defensible, and that he is not contradicting his own moral theory when he argues that we have a genuinely *moral* duty to promote the highest good, which includes happiness. But before presenting the full defense, it is important to see just why Kant seems to be in danger of self-contradiction.

Let us recall that, for Kant, the "highest good" refers to the proper and complete *object* of the will, which we have a moral obligation to bring about. But secondly, Kant *also* seems to claim that, because of the utter *formality* of the moral law, a will that adheres to the moral law can really only have *itself* (understood as legislating universal law) as its proper *object*.

This second line of reasoning is present in Kant's *Grounding of the Metaphysics of Morals*. There Kant argues that the moral law, as pure and apriori, cannot be identified with any particular, materially-determinate positive law; therefore, the content of the moral law can be nothing other than the *form* of lawfulness as such, or universality.[16] This gives rise to the first formulation of the categorical imperative: "Act only according to that maxim whereby you can at the same time will that it should become a universal law" (*GMM*, 4:421). Accordingly, an action that is willed out of duty is morally right to will, not because of the external object brought about by the will, but rather because it is the kind of action whose maxim can be willed as a universal law. It is the kind of action whose maxim could also be willed by every other rational being, and thus it is a maxim that I—insofar as my willing of it may affect others—do not need to *impose* on the free willing of others. To the extent that I violate the universalizability-test of the first formulation, I also implicitly impose my will on others, and thus use them as a means to my own ends. Thus we have the second formulation of the categorical imperative: "Act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means" (*GMM*, 4:429). These two formulations, when considered together, yield Kant's third formulation, the so-called formulation of the "kingdom of ends," which involves a systematic union of rational beings wherein each legislates universal law for oneself while at the same time being subject, like all the rest, to such law (*GMM*, 4:433).

Notice that the three formulations of the categorical imperative are silent about the specific *object* of our moral willing, other than to say that this *object* must somehow involve the very act of moral willing itself. The reason for the silence is easy to discern: because the moral law must be grounded on autonomous, self-legislating reason alone, it cannot have an object whose content is derived from any external (heteronomous) source. Thus the moral law must be utterly formal, "one which prescribes to reason nothing more than the form of its giving universal law as the supreme condition of its maxims" (*CPrR*, 5:64). But this formality, in turn, implies that the will's moral duties have nothing to do with external objects or consequences, but only with the relation of the will to itself: ". . . here the concern is with objectively practical laws, and hence with the relation of a will to itself insofar as it is determined solely by reason. In this case everything related to what is empirical falls away of itself . . ." (*GMM*, 4:427). Accordingly, a will

that acts out of respect for the moral law "will therefore be *indeterminate as regards all objects* and will contain merely the form of willing; and indeed that form is autonomy" (*GMM*, 4: 444; emphasis added). Finally, as Kant writes, the moral law commands only "that everything be done from the maxim of such a will as could at the same time *have as its object only itself* regarded as legislating universal law" (*GMM*, 4:432; emphasis added).

The implications to be drawn from these statements seem to be devastating for Kant's notion of the highest good. The highest good, insofar as it includes happiness, seems to be a *determinate* object, whose content apparently comes from an *external* (empirical) source. Furthermore, the alleged duty to promote the highest good seems to imply that the will must have as its object something much more determinate and external than "only itself regarded as legislating universal law." Along these lines, Lewis White Beck (one of this century's most influential commentators on the second critique) concludes that the notion of the highest good ultimately contradicts the rest of Kant's moral theory:

> . . . if the 'possibility of the highest good' means anything more than its necessary condition [namely, the will's adherence to the moral law], it is to that extent incompatible with what [Kant] has said earlier and more consistently about the lawful form of the maxim itself being the object of the moral will.[17]

To conclude: if Kant's account of the highest good involves the kind of heteronomy and externality indicated above, then it seems that we can have no properly *moral* obligation to promote the highest good (as something beyond the will itself, regarded as legislating universal law); but if that is the case, then Kant's moral proof ultimately fails.

4. Defense of the Moral Proof

Perhaps the best way to defend Kant is to begin by elucidating what may seem rather obvious: the condition of the possibility of my having any moral duty at all is that, *prior to* and *apart from* any act of willing on my part, there must be some pre-existing state of affairs that *ought not to be* as it is, i.e., there must be some already-actual, morally deficient, state of affairs that *ought* to be made *different*. For if there were no such state of affairs *prior to* and *apart from* my act of willing, I would have no moral duty and thus no morally significant reason to will anything at all.[18] Now since the morally deficient state of affairs that grounds and makes possible a moral "ought" for me must pre-exist in the world itself, apart from whether or not I undertake to will anything at all, my moral duty must not merely be to bring about my own willing alone, but also to bring about a change or transformation in the pre-existing and morally deficient world that gives rise the "ought" in the

first place. In other words, my moral obligation must be to bring about some actual result, that is, an object *other* than just my own act of willing, even regarded as legislating universal law.

Thus I can never have a *moral* obligation to bring about *only* my own act of willing; the proper object of my willing can never be just itself regarded as legislating universal law.[19] This is why Kant never says that our will *actually* ought to have as its object "only itself regarded as legislating universal law." What he says, rather, is subjunctive and counterfactual: the moral law commands "that everything be done from the maxim of such a will *as could* at the same time have as its object only itself regarded as legislating universal law" (*GMM*, 4:432; emphasis added). In fulfilling one's moral duty, one ought to act *as if* the sole object of one's will were simply one's own act of willing, regarded as legislating universal law; but in order to be confronted with a moral duty in the first place, the object to be realized as a result of one's willing must be *other* than just the act of moral willing alone. The fact that one has a moral obligation to promote an object that is *other* than or *beyond* the mere act of moral willing, is a necessary *structural* feature of *moral obligation* itself, and Kant rightly realized this. It is not, as Beck claims, the result of an illicit appeal to concerns that are external to morality and thus heteronomous.[20]

Now if, in order to have a moral obligation at all, the moral obligation must be to bring about some object that is *other* than just my own act of willing, what kind of object might this be? Can anything more be said about the *object* that the good will ought to bring about as *other* than itself? Indeed, much more can be said.

Even though the object of one's moral willing must be *other* than the act of moral willing alone (regarded as legislating universal law), the goodness of the object to be willed cannot simply be "good in itself," entirely independent of the good will's willing of it; otherwise, the will's autonomy would be compromised (*CPrR*, 5:109). This, for Kant, is due to the fundamental paradox involved in any critical examination of practical reason: the morally good will does not will what it wills because the object is antecedently good; rather, the object being willed is morally good insofar as it can be willed in accordance with the maxims of a morally good will.[21]

So our problem is now the following: what kind of object, if any, do we have a duty to promote that is genuinely *other* than the mere act of moral willing (regarded as legislating universal law), yet not good in itself, entirely independent of such willing? Such an object would have to be a product of the will's act of willing, and thus a *result* of some kind, but a result that would be good, not in itself, but only to the extent that it is related to the will's willing of it. Such an object could not simply be *the fact of* the will's willing it; for such a will, after all, could be completely ineffectual and

produce no object other than itself. Rather, such an object would have to be the fact that some object *other* than the act of willing was *actually produced in the world* just as the will willed it—provided, of course, that the will's willing of it were consistent with the moral law itself. In short, the object would have to be (at the very least) the *fact that* things *in the world* do, indeed, turn out as one wishes and wills—provided, of course, that such was consistent with the moral law.

Now this state of affairs where things in the world do turn out according to "wish and will" is simply the *formal* and *indeterminate* definition that Kant gives to his notion of "happiness" (*CPrR*, 5:124; see also *CPR*, A806/B834); as we have seen, any *material* or *determinate* definition would introduce heteronomy into our understanding of the will's necessary object. Now we are in a position to see that this *external*, yet only *formally*-articulated object that we have been describing—namely, the fact that things in the world do turn out according to wish and will, in proportion to one's adherence to the moral law—is nothing other than what Kant has been calling the "highest good." This object does indeed pertain to *results* that are genuinely *other* than the mere act of moral willing; however, the goodness of such results are not good in themselves entirely independent of the good will's willing of them. Rather, they are good only conditionally and in relation to such willing.

With this, we can now see how Kant can consistently hold that the *determining ground* of the good will is nothing but the good will itself, regarded as legislating universal law; but that the *object* of a completely good will must be something *external to* and *other than* just the good will's act of willing. Accordingly, Kant can say without detriment to his autonomy-based moral theory that we do have a genuinely *moral* obligation to promote the highest good as the *object* of our willing. While the happiness that is a component in this highest good is an object that is genuinely *other* than one's own moral willing, this happiness is only *formally* defined and thus entails no heteronomy in the will's determining ground. Furthermore, because the highest good pertains to results beyond the mere act of our moral willing, it is an object whose realization in the world is not entirely within the control of our agency; that is, it is an object whose realizability in the world cannot be conceived without the postulate of God.

Kant's doctrine of the highest good does not involve any particular notion of happiness as a special kind of pleasure or contentment; rather, it rests simply on the notion that the moral or virtuous will should not be frustrated or prevented from bringing about in the world that which it wills, insofar as such willing is consistent with the moral law. Thus the harmony of virtue and happiness in the highest good is willed *from the side* of virtue and *for the sake* of virtue. The resulting happiness is good, not because it is

good in itself, but because it is good that virtuous wills not be frustrated in this world. This account also shows more specifically what the highest good means with respect to justice. It is unjust, from a purely moral point of view, that the wills of those who are moral and virtuous should be frustrated, or that the wills of those who are immoral and vicious should succeed, in this world. Just as we have a moral duty to promote the highest good in the world, so too do we have a duty to promote justice—and the conceivability of achieving such justice requires the postulate of God.

5. Implications of the Moral Proof: Circularity and Self-referentiality

This analysis and defense of Kant's moral proof reveals an extremely interesting feature about what it means for a finite will to be subject to moral obligation.[22] As we have seen, the condition of the possibility of being subject to moral obligation is that the *object* of one's will, the product or state of affairs that one ought to bring about as a result of one's willing, must be *other* than one's mere act of moral willing. As we have also seen, the complete object of the morally good will is the highest good, understood as virtue combined with proportionate happiness. But precisely because this object to be promoted must be something *other* than the finite will's own moral willing, the morally good finite will cannot produce its complete object entirely out of itself alone; it depends also on the prior givenness of the natural world. The fact that one has a moral duty to change the world at all presupposes that there is a pre-existing, morally deficient world to be changed. Thus there must be some pre-given nature (or "raw material" for morality) out of which the finite will must seek to bring about its proper and complete object, the highest good. Furthermore, since this nature is at first simply given and not already a result of the finite rational being's own moral willing, it must at first appear to be indifferent and unrelated to the finite rational being's moral purposes. And as we have seen, it is for this very reason that it seemed impossible for us to bring about the highest good on our own; and thus it became necessary to believe in God as a practical postulate.

But these reflections have now led us into a circle. The fact that the morally obligatory object of our moral agency must be *other* than the our moral willing alone (regarded as legislating universal law) explains not only why we *have* the obligation in the first place, but also why it seems *impossible* for us to fulfil the obligation on our own, without the postulate of God's existence. On the one side, it is possible for us to have a moral obligation in the first place only because a pre-existing and morally deficient world is already *given* to us, apart from any act of moral willing on our part; it is precisely such givenness that precedes and elicits our properly moral willing

in the first place. But conversely, it seems impossible for us to fulfil any such moral obligation on our own (apart from the postulate of God's existence) precisely because this pre-existing, morally deficient state of affairs is itself not already the result of our moral activity, and thus must first appear to be wholly indifferent and unrelated to our moral purposes. That is, the reason why we can have a *moral obligation to promote the highest good* as an object in the first place is the *very same reason* why we *cannot think of ourselves as actually promoting it without the postulate of God's existence.* In more Kantian terms: the condition of the *possibility* of our having any moral obligation at all (where this condition is understood as the givenness of a world that initially confronts us as not already the result of any properly moral activity) is at the same time the condition of the *impossibility* of our doing our moral duty, except on the supposition that God exists (i.e., except on the supposition that the given world already *is* the result of some moral activity, namely God's).

Thus, in a sense, Kant's moral argument *is* indeed circular. Our moral obligation seems impossible for us (and thus leads us to the postulate of God's existence) for the very same reason that we take ourselves to have the obligation in the first place: namely, because the object that we are morally obliged to produce must be *beyond* our mere act of moral willing and thus cannot be produced entirely *out of* our moral willing alone, but must instead be produced out of a pre-existing nature that is *given* in advance of our moral activity and thus seems to be indifferent and unrelated to our moral purposes. One can escape the circle and thus deny the force of the argument, if one calls into question the claim that we have such a moral obligation in the first place. But can one really question this claim without implicitly contradicting oneself? If one cannot do so, then it would be self-refuting to question this claim, in which case Kant's argument would be implicitly self-validating. And if that is the case, then the argument's circularity or self-referentiality—far from being a *vicious* circularity—is actually one of the argument's strengths.

In what follows, I shall argue that we do, indeed, have a moral obligation to promote the highest good (formally defined), and that there is a kind of self-validating certainty in our awareness of such an obligation. It is thus immediately self-refuting to doubt that one has this obligation, and mediately self-refuting to doubt the postulate of God's existence (insofar as belief in God's existence is the ground of one's hope of fulfilling the obligation). If my argument works, then Kant's moral proof would involve a self-validating self-referentiality (or circularity) that overcomes the uncertainty that typically affects theoretical proofs of God's existence. After all, theoretical proofs typically rely on evidence that is contingently and externally given to consciousness by a source that is other than consciousness itself; thus they tend to rely on evidence whose evidentiary character can always be doubted without

self-contradiction. By contrast, the evidentiary character of the evidence that is operative in Kant's moral proof is not located in any *external* data *as given*, but rather in one's incorrigible, inner awareness of having a moral obligation. That is, the evidence is simply one's own *activity* of knowing that one *ought* to make the world *different* than it currently is. But how can one's awareness of this moral obligation be shown to be self-validating?

We have already seen that our obligation to bring about the highest good must be understood as *formal* and *indeterminate*. The obligation to bring about the highest good is simply the obligation to bring about a state of affairs in which things in the world do turn out as the good will wills them; it is the obligation to bring about a harmony between morality (what *ought* to be the case) and nature (what *actually is* the case). Now as we have also seen, the condition of the possibility of being conscious of such an obligation is that one is confronted by a morally deficient world that *ought* to be made different than it actually *is*, that one is conscious of a discrepancy between what *is* and what *ought* to be. (It is precisely this consciousness that grounds our awareness of both *having* a moral obligation and *being unable to fulfil* our obligation, except on the supposition that God exists.) But how is this consciousness of having a moral obligation (this consciousness of a discrepancy between what *is* and what *ought* to be) self-validating? If anyone questions whether there is such an obligation, the very act of questioning implicitly demonstrates the existence of the obligation at issue. For every genuine question as such is a *challenge* to the current state of the world *as given* to the questioner. Every question reveals the questioner's awareness that the world as it *is*, is not exactly as it *ought* to be. Even a "purely theoretical" question presupposes this discrepancy between what is and what ought to be, since even a theoretical question seeks an answer that does not already exist for the questioner. Thus even a purely theoretical question presupposes that a world in which I do not have an answer is *deficient* in some minimal sense, and *ought* to be changed.

Thus the condition of the possibility of questioning is the questioner's own immediate awareness of some discrepancy between what *is* and what *ought* to be. Correspondingly, a person's activity in questioning demonstrates his or her immediate awareness of an obligation to bring about change in the world as given. Thus the activity of questioning self-referentially validates one's awareness of being obliged to bring about change in the world, which in turn leads to Kant's moral proof. For if I am obliged to bring about change in the world (even if this change involves nothing more than overcoming my current state of ignorance), I can genuinely hope to achieve this change only on the supposition that my moral purposes, and the given world to be transformed by them, have a common ground in God.

But isn't it still possible for someone to escape this argument by disputing the claim that one really has an *obligation* to ask questions? After all, how do I really know that I *ought* to ask questions? The answer, once again, must be given by way of self-reference. I cannot even begin to question my obligation to ask questions except by undertaking to fulfil that obligation itself. If I am to be fully self-consistent in questioning whether I ought to ask questions, then my only course of action—like that of Aristotle's skeptic—is to remain silent. However, my silence cannot be a silence that I *arrive at* as a result of asking questions and realizing that consistency requires silence. Rather, my silence must be an animalistic silence that *precedes* and is entirely *separate from* all question-asking; otherwise, even my silence is not fully self-consistent. But a purely animalistic silence is impossible for me.

Of course, one may object to this account by saying that the virtual identification of our *obligation to ask questions* with our *obligation to transform the world* (however minimally) tends to blur the distinction between *theoretical* reason and *practical* reason. After all, it seems that theoretical reason has to do with asking questions and *knowing* the world, while practical reason has to do with undertaking action and *transforming* the world. But as I have been suggesting here, question-asking is itself ultimately practical and aims at transforming the world, insofar as my current state of ignorance is part of the world. Kant himself makes the same point, but with much broader implications, when he tells us that "all interest [of reason] is ultimately practical and even that of speculative reason is only conditional and is complete in practical use alone" (*CPrR*, 5:121).[23]

Summarizing the argument: to the extent that I always already *find* myself to be a questioner, I am inescapably committed to the position that I *ought* to ask questions. If I want to escape the obligation of asking questions, I always arrive on the scene too late, for my finding of myself as a questioner is itself always the *result* of the very question-asking that I would seek to escape. Now as a questioner, I am necessarily committed to the position that the world as given to me—including my current state of ignorance—is not as it *ought* to be. Furthermore—as we have seen above—the world first confronts me as not being the result of any properly moral activity, and thus appears as indifferent and unrelated to my moral purposes. Thus I can hope to fulfil my obligation to transform the world as given only on the supposition that God exists.

The undeniable fact of my inescapable obligation to ask questions, then, demonstrates the justifiability of my belief in God. In this respect, Kant's moral proof is oddly similar to the *ontological proof*, since it rests not on any external evidence, but on the immediate and incorrigible evidence of what I can or cannot conceive without self-contradiction. In the present case, I simply cannot—without self-contradiction—question whether I ought

to ask questions. Of course, I might never have begun asking questions in the first place; but now that I have found myself as a questioner, I am necessarily committed to the position that I ought to ask questions. Furthermore, I need to believe that the world as it is given to me—including my current state of ignorance—can be changed as a result of my agency; and I need the postulate of God in order to believe that such change is possible. Thus my *assertiveness* in challenging the world through my question-asking must always be combined with the *modesty* of knowing that I can hope to transform the world only on the supposition that an all-powerful, all-benevolent God exists. As Kant writes, the *self-esteem* that human beings achieve through their observance of a self-legislated moral law must always be combined with *humility* (*CPrR*, 5:128).[24]

Michael Baur, Fordham University

Notes

1. Lewis White Beck, for example, writes that "Kant's usual high-quality workmanship is not much in evidence" in his "moral proof" and discussion of the "highest good." See Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason* (Chicago: University of Chicago Press, 1960), p. 246. See also Keith Ward, *The Development of Kant's View of Ethics* (Oxford: Basil Blackwell, 1975), p. 85; and Thomas Auxter, *Kant's Moral Teleology* (Macon, Ga.: Mercer University Press, 1982), pp. 83–85.

2. See Immanuel Kant, *Critique of Pure Reason*, trans. by Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1997), pp. 551–589. All subsequent references to the *Critique of Pure Reason* will be based on this translation and cited parenthetically in the text in the following format: *CPR*, A567/B595-A642/B670 (referring to the pagination in the 1781 A-edition and 1787 B-edition).

3. Immanuel Kant, *Critique of Practical Reason*, in *Immanuel Kant: Practical Philosophy*, ed. and trans. by Mary Gregor (Cambridge: Cambridge University Press, 1999), p. 258. All subsequent references to the *Critique of Practical Reason* will be based on this translation and cited parenthetically in the text in the following format: *CPrR*, 5:147 (referring to the Akademie-edition volume and page number).

4. This is a quotation from Kant himself (*CPrP*, 5:144), referring to Wizenmann's work published in the *Deutsches Museum* (February, 1787). For more on the thought of Thomas Wizenmann, see Frederick Beiser, *The Fate of Reason* (Cambridge: Harvard University Press, 1987). pp. 109–113.

5. For more on this, see Christine M. Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), p. 28.

6. Garve had articulated his criticism in his "Versuche über verschiedene Gegenstände aus der Moral und Literatur." Kant's response to this criticism appears in his essay, "On the Common Saying: That may be correct in theory, but it is of no use in practice." See *Immanuel Kant: Practical Philosophy*, ed. and trans. by Mary Gregor (Cambridge: Cambridge University Press, 1999), pp. 277–309. All subsequent references to the *Theory and Practice* essay will be based on this translation and cited parenthetically in the text in the following format: *TP*, 8:277–309 (referring to the Akademie-edition volume and page number).

7. See *TP*, 8:281–289.

8. Now what, specifically, does Kant mean by this proportionality or harmony between virtue and happiness, freedom and nature, in the highest good? At least two different interpretations are possible. Under one interpretation, "the highest good" means simply a connection and proportionality between virtue and happiness, even if the optimal amount of virtue (and thus happiness) is not actually realized. Under this interpretation, the "highest good" is realized if those who are virtuous are happy in proportion to their virtue, and those who lack virtue also lack happiness, again in proportion to their lack of virtue, so that justice is done. This interpretation is supported by Kant's statements at *CPrR*, 5:114, and *CPR*, A812/B840. Under a second interpretation, "the highest good" means not only a proportionality between virtue and happiness, but also the complete realization of both, such that everyone is both virtuous and happy, and justice is thereby done. This interpretation is supported by Kant's statements in section 88 of the *Critique of Judgment*. See Immanuel Kant, *Critique of Judgment*, trans. by Werner S. Pluhar (Indianapolis: Hackett Publishing Company, 1987), p. 343. All subsequent references to the *Critique of Judgment* will be based on this translation and cited parenthetically in the text in the following format: *CJ*, 5:453 (referring to the Akademie-edition volume and page number).

In light of all the relevant textual evidence, it seems that Kant's own understanding of the highest good is best expressed by the second interpretation. For if, as Kant says, we have a moral obligation to promote optimal virtue in ourselves and others, and if the highest good entails a proportionality between virtue and happiness, it seems to follow that our duty to promote optimal virtue as well as the highest good also entails a duty to promote optimal happiness (as proportionate to such optimal virtue). That is, we have a duty to promote not only a proportionality between virtue and happiness, but also the complete realization of both, such that everyone is both virtuous and happy. Under this interpretation, the highest good is a Kantian kingdom of ends *and* of nature, wherein each person treats all persons (including oneself) as ends in themselves, and is also fully happy in doing so.

For more on the two different interpretations, see John E. Hare, *The Moral Gap: Kantian Ethics, Human Limits, and God's Assistance* (Oxford: The Clarendon Press, 1996), especially pp. 71–74.

9. See Kant's critique of Epicureanism and Stoicism: both schools of thought tried to overcome the difficulty by arguing—in different, but equally erroneous ways—that the connection between virtue and happiness was merely analytic (*CPrR*, 5:126–127).

10. As Kant writes, "Morals is of itself practical in the objective sense, as the sum of laws commanding unconditionally, in accordance with which we *ought* to act, and it is patently absurd, having granted this concept of duty its authority, to want to say that one nevertheless *cannot* do it. For in that case this concept would of itself drop out of morals (*ultra posse nemo obligatur* [*no one is obligated beyond what he is able to do*]) . . . ." From Immanuel Kant, "Toward Perpetual Peace: A Philosophical Project," in *Immanuel Kant: Practical Philosophy*, ed. and trans. by Mary Gregor (Cambridge: Cambridge University Press, 1999), p. 338. The Akademie-edition reference is 8:338 (referring to the volume and page number).

11. Thus, in a sense, Kant's argument about the highest good and the divine will is similar to his discussion of teleology in the *Critique of Judgment*. In the *Critique of Judgment*, Kant argues that we are justified in thinking in terms of teleology (as a regulative idea) in order to account for the "wholly fortuitous" and otherwise inexplicable connection between our theoretical reason and the manifold givenness of nature (*CJ*, 5:406–407).

12. See also *CPrR*, 5:140, where Kant writes: "This Being must be omniscient, in order to be able to know my conduct even to the most intimate parts of my intentions in all possible cases and in the entire future. In order to allot fitting consequences to it, he must be omnipotent, and similarly omnipresent, eternal, etc. Thus the moral law, by the concept of the highest good as the object of a pure practical reason, defines the concept of the First Being as that of a Supreme Being . . . . Therefore, the concept of God is one which belongs originally not to physics, i.e., speculative reason, but to morals."

13. Thus Kant writes: "So far as practical reason has the right to lead us [and it does], we will not hold actions to be obligatory because they are God's commands, but will rather regard them as divine commands because we are internally obligated to them" (*CPR*, A819/B847). By the same token, no being—not even God—has a right to make use of a human being as a mere means to an end (*CPrR*, 5:131). For more on this, see John R. Silber, "The Importance of the Highest Good in Kant's Ethics," *Ethics* 73 (April, 1963): 179–197; see especially p. 189.

14. See, for example, Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason*, p. 246; see also Theodore M. Greene, "Introduction" to Immanuel Kant's *Religion Within the Limits of Reason Alone*, trans. and ed. by Theodore M. Greene and Hoyt H. Hudson (New York: Harper and Row, 1960), p. lxiv.

15. See, for example, *CPrR*, 5:113: "It is apriori (morally) necessary *to produce the highest good through the freedom of the will.*"

16. "But what kind of law can that be the representation of which must determine the will, even without regard for the effect expected from it, in order for the will to be called good absolutely and without limitation? Since I have deprived the will of every impulse that could arise for it from obeying some law, nothing is left but the conformity of actions as such with universal law, which alone is to serve the will as its principle. . . ." Immanuel Kant, *Groundwork of the Metaphysics of Morals,* in *Immanuel Kant: Practical Philosophy*, ed. and trans. by Mary Gregor (Cambridge:

Cambridge University Press, 1999), p. 56. All subsequent references to the *Groundwork* will be based on this translation and cited parenthetically in the text in the following format: *GMM*, 4:402 (referring to the Akademie-edition volume and page number).

17. See Lewis White Beck, *Commentary on Kant's Critique of Practical Reason*, p. 243.

18. In fact, if I *did* undertake some action to make the world different when, in fact, there was no moral deficiency in it at all, then my instigation of such an action could itself be morally wrong.

19. Silber makes a similar observation in his famous article on Kant's notion of the highest good. Once Kant has finished talking about the formal, moral law as the determining ground of the will, Silber writes, he has "gone as far as he can from the side of the law alone in the determination of a material object of volition. He now confronts an ethical paradox . . . , namely the paradox of the willing of the willing of nothing." See Silber, "The Importance of the Highest Good in Kant's Ethics," p. 190.

20. See Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason*, pp. 243–244.

21. ". . . the concept of good and evil is not defined prior to the moral law, to which, it would seem, the former would have to serve as a foundation; rather the concept of good and evil must be defined after and by means of the law" (*CPrR*, 5:63).

22. As Kant notes, it is impossible for an infinite will to be subject to moral obligation. See *GMM*, 4:414.

23. For an account of the highest good as the ultimate end of both practical and theoretical reason, see Mary-Barbara Zeldin, "The Summum Bonum, the Moral Law, and the Existence of God," *Kant-Studien*, vol. 62 (1971): 43–54.

24. One of Kant's earliest proponents, J. G. Fichte, poignantly describes this unity of self-esteem and humility: "The first person who asked a question about the existence of God broke through the boundaries; he shook mankind to its deepest foundations and brought man into a conflict with himself which has not yet been resolved and which can only be resolved by proceeding boldly to that supreme point from which the practical and the speculative appear as one. Presumption led us to philosophize, and this cost us our innocence. We caught sight of our nakedness, and since then we have had to philosophize for our own salvation." From Fichte's letter to Jacobi, August 30, 1795. See *Fichte: Early Philosophical Writings*, trans. and ed. by Daniel Breazeale (Ithaca: Cornell University Press, 1988), p. 412.