

The Right vs the Good: Kant vs Rawls

[*O Justo vs o Bem: Kant vs Rawls*]

Patricia Kitcher¹

Columbia University (New York, USA)

DOI: 10.5380/sk.v%vi%i.92097

Abstract

Rawls regarded the priority of the right over the good as the characteristic feature of Kantian constructivism. I have four goals in the paper. First, I try to refute Rawls's reading of Kant on the relation between the right and the good. Second, I fill out Kant's picture of the rational natures that have intrinsic value: They have the law of duty within, are predisposed to respect themselves and others who have the law of duty within, and belong to a community of rational natures. Third, I argue that because Kant thought that the right and the good were coeval, he is not a constructivist, but a kind of realist. Finally, I use my examination of the good, the intrinsically valuable in Kant, to reject any temptation to regard his ethics as dependent on his teleological claim that human nature is the end of nature.

Keywords: Rawls; right; good; constructivism; motivating ground; moral community.

¹ Patricia Kitcher is the Roberta and William Campbell Professor Emerita of the Humanities, Professor Emerita of Philosophy, Columbia University, FAAS. E-mail: pk206@columbia.edu

“This priority of the right over the good is characteristic of Kantian constructivism.”
(Rawls, 1980/1999, 319)

“As Kant puts it ... ‘The paradox is that 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.’”
(Rawls, 2000, 227)

1. Rawls and Kant

By carefully using “Kantian” to describe his work, Rawls tried to forestall identifying his attractive and influential Kant-inspired theory as Kant’s own. On the issue of the priority of the right over the good, however, Rawls seemed to think that there was no daylight between their positions. After citing Kant’s observation about the paradox of not defining good and evil before the moral law, he notes that emphasis is added to the whole sentence, thereby suggesting that Kant not only held this view, but that he regarded it as a central feature of his ethics – as it is of Rawls’s ethical constructivism. I will argue that Rawls misreads the paradox passage and that the doctrine that he attributes to Kant is inconsistent with Kant’s view of the relation between the first and second formulations of the Categorical Imperative. Where I agree with Rawls is on the centrality of the relation between the right and the good in an ethical theory to the nature of the theory. Rawls took Kant to prioritize the right over the good and so to be some type of moral constructivist (*avant la lettre*); I take Kant to regard the right and the good as necessarily coeval and so to be a kind of moral realist. My goal in pursuing the relation between the right and the good in Kant’s theory is less to refute Rawls’s (now) canonical reading – though I would like to do that – than use his issue to illuminate the status of Kant’s moral theory.

2. The Evidence Against Kant Prioritizing the Right over the Good from *Groundwork 2*

After presenting his discovery of the supreme moral principle as the Categorical Imperative (*Act only on that maxim that you can also will to become a universal law* [GMS, AA 04: 421]), Kant raises an odd question:

Is it a necessary law for all rational beings always to judge actions according to maxims which they can will to be universal laws? (GMS, AA 04: 426)²

His question is neither rhetorical, nor skeptical, but substantive. He takes himself to have already shown in *Groundwork 1* that morality, as it is ordinarily understood, requires that agents be able to act for solely moral reasons. Since he has just shown that the Categorical Imperative is the supreme moral principle, it follows that morality will be possible only if the world contains beings of a certain kind, beings who, not through or with the aid of any sensory incentive, but

² References to Kant’s works other than the *Critique of Pure Reason* will be given in the text by the volume and page of Kant (1900-). References to the *Critique of Pure Reason* are in the text with the usual A/B pagination. Translations for the *Groundwork of the Metaphysics of Morals* are from Gregor and Timmermann (2011); translations for the *Metaphysics of Morals* are from Gregor (1996), translations for the *Critique of Practical Reason* are from Pluhar 2002; translations for the *Critique of Pure Reason* are from Pluhar (1996). In all cases, however, I indicate Kant’s emphases through bold not italics and I translate “*Vorstellung*” and related forms as “representation,” not Pluhar’s by preferred “presentation.”

solely through their rational nature, always evaluate actions through the Categorical Imperative. Hence, the importance of asking his odd question. Kant laments that his question can be answered only by taking a reluctant step beyond our understanding of morality into metaphysics (GMS, AA 04: 428).

Kant does not answer the question of the existence of such creatures immediately, but turns to moral psychology, noting that a will must always have an end. If we look back to *Groundwork 1*, we can bring the problem he is addressing into sharper focus. There Kant presents two versions of a case of someone in need. In the first version, the person helps the needy individual, because he takes pleasure in helping others and making them happy. In the second, the same person has lost his ability to take joy in the happiness of others, but still has the resources to help the needy individual and does so, because he recognizes that it is his duty to help. In both versions, the purpose or end of the action is to help the needy individual for his own sake. Something else is involved in the second case, however, namely, acting from duty. I take Kant's question as he moves from the first to the second formulation of the Categorical Imperative to be: What end can a will have in relieving the need for the right reason, out of respect for moral law?

Kant explains that what is required for the possibility of morality is an objective and motivating ground or basis – *Grund* – for the will. I have tried to make his inquiry more intuitive by using one of his examples. The question of the ground of respect for the moral law can also be raised in a general way. Morality is possible only if agents can adhere to the results of the first formulation of the Categorical Imperative, henceforth, 'CI1', tests. But what is the agent's basis, ground, motive for refraining from an action if it fails one of the tests? Why should she restrict what she does to actions that she could will that everyone performs? Kant's implicit question is, again, not skeptical, but substantive. He provides parameters for the answer by contrasting the moral case with the incentives that motivate the pursuit of happiness. In the latter case, the grounds for action vary from person to person, depending on what makes them happy. For CI1 to be binding on agents, however, its motivating ground cannot vary, cannot depend on the particular or subjective ends of individuals. It must be invariant across all moral agents, or objective. Here, as in the *Critique of Pure Reason*, Kant uses 'objective' to indicate both invariance across subjects and relation to an object (e.g., B 138).

Kant juxtaposes his implicit question about the motivating ground for following CI1 and his odd question about whether it is a necessary law for rational beings to judge their actions according to CI1, because they have the same answer: the existence of agents with the moral capacity that he has delineated in *Groundwork 1 and 2*, the capacity to act out of respect for moral law. Kant abbreviates this capacity to that of having a "rational nature" (GMS, AA 04: 429), because he anticipates the positive answer to his odd question that is supposed to come in *Groundwork 3*. The law of duty lies *a priori* in reason, so those with a rational nature necessarily judge their actions according to maxims that they can will to be universal laws.

The most crucial textual fact about the second formulation of the Categorical Imperative (CI2)

Act in such a way that you treat humanity, whether in your own person or anyone else's, never merely as a means, but always also as an end. (GMS, AA 04: 429)

is that Kant tells the reader explicitly in a footnote, and implicitly through his language, that he cannot argue for it as a formulation of the supreme moral principle until *Groundwork 3*. Why does Kant present CI2 in *Groundwork 2*, before he is able to argue for it? He seems to answer this question at the end of the section when he explains that both the first two sections are analytic – are explicating the ordinary concept of morality (GMS, AA 04: 445). Presumably he thinks that CI2 and the third formulation, CI3, which he also presents in *Groundwork 2*, are important aspects of ordinary moral understanding and must be made explicit to have a tolerably complete analysis.

We can see why Kant might think that completeness requires CI2 through considering the relation between CI1 and CI2. It is important to bring in CI2, because it gives the basis or motivating ground for agents to follow CI1. On the other hand, CI1 clarifies what CI2 requires. How you show your respect for other persons is by not arrogating rights to yourself that you would not grant to all others. With this fuller picture, we can understand why Kant maintains that the different looking principles are equivalent, in the sense that it does not matter which an agent uses to figure out what she may or ought to do. Consider again the case of helping a needy individual from the motive of duty. It does not matter whether the agent foregrounds CI1 or CI2. She is either respecting humanity in her own person and that of others, by only engaging in actions that she could will that everyone do or she is not allowing herself to do something that she could not will others to do, out of respect for humanity. In the case of morally worthy actions, both elements need to be involved.

Given the relations between CI1 and CI2, we can already see that there is a strong textual case against taking Kant to prioritize the right over the good. Without the good, the existence of creatures with the rational moral capacity, there would be no suitable motivating ground for following the law of duty, and without the law, there would be no adequate means of expressing respect for all rational natures. Kant's summary of the relations among the formulae says exactly this: any maxim must have both a form and a matter (GMS, AA 04: 436). Specifically, any morally worthy maxim must have the form of universalizability and must have rational nature as its matter – as the end to which the action is directed. Since both the form of law and the end of rational nature must be present in morally worthy action, they must be coeval. Since the form cannot be prior to the matter, the right cannot be prior to the good.

What about the textual evidence on the other side? It seems to me that it is weak. Both at the beginning of the section of the *Critique of Practical Reason* where Kant discusses good and evil and right after his comment about a paradox, he makes the same point:

If the concept of the good is not to be derived from any antecedent practical law, but is rather to serve as its basis, then it can only be the concept of something whose existence produces pleasure and thus determines the causality of the subject to produce this something. (KpV, AA 05: 58, see also 05: 63, my italics).

Rational nature is not, however, a good of this type. It is an existing end that all agents have, not an end that different agents acquire through their varying experiences with pleasure and then try to bring into being for the pleasure that it will yield.

3. The Evidence Against Kant Prioritizing the Right over the Good from his Theory of Predispositions

Although Kant argued that morally good action requires both the form of universal law and the motivating ground of respect for rational natures, he might have taken the right to be prior to the good in a different sense. Kant memorably claims at the end of the *Practical Critique* that one of the two things that fill any human mind with admiration and reverence is “**the moral law within me**” (KpV, AA 05: 161, my italics). Earlier, in §7, he had argued that

So act that the maxim of your will could always hold at the same time as a principle of universal legislation. (KpV, AA 05: 30)

is a “basic law of pure practical reason” (KpV, AA 05: 30). And in a note to *Religion within the Bounds of Bare Reason*, he maintains that

If this law [CI1] were not given *within* us, no reason would ever enable us to cogitate it as a law or to talk the power of choice into it. (RGV, AA 06: 26a, my italics).

I understand all these texts to say the same thing: As the categories lie *a priori* in the mind waiting to be awakened to action by the receipt of sensory information, so too, the Categorical Imperative lies *a priori* in the mind waiting to be awakened into action by the agent's need to decide on a course of action (see also *RGV*, AA 06: 29). It is a basic principle of pure practical reason, a principle that pure reason uses to figure out which actions are permitted or required. The *Religion* book characterizes this law given to the mind as a "predisposition" to "personality." It is a necessary condition for the possibility of "personality," of being an agent with the moral capacity (*RGV*, AA 06: 27). As a predisposition, it can be neither acquired nor extirpated (*RGV*, AA 06: 28). It might seem that the presence of the law of duty in human reason is necessarily prior to respecting creatures with the moral capacity. After all, they must have the law within, before they can be respected by virtue of having it.

If we turn to the *Metaphysics of Morals*, however, we find more predispositions for morality (*MS*, AA 06: 399). Of relevance to the point at issue, humans must have a predisposition – not a duty – of self-esteem.

It is not correct to say that a man has a **duty of self-esteem**; it must rather be said that the law within him unavoidably forces from him **respect** for his own being ... It cannot be said that a man **has** a duty of respect toward himself, for he must have respect for the law within himself in order even to think of any duty whatsoever. (*MS*, AA 06: 402-403)

Even without working through this passage in the detail it requires, Kant is clearly making two points. First, those who have the law of duty within are thereby forced to respect their own beings. Respect for yourself as a moral agent is no more voluntary than having the law within is voluntary. Second, it makes no sense to talk about the law as prescribing duties and then to ask whether it also needs to be respected. Unless it is respected, it cannot prescribe any duties. What I leave unclear for the purposes of this paper is whether the moral law can be respected only through respecting creatures who have it within. One possible counterexample would be angels, who have the law within, but who lack opposing sensible desires and so automatically follow the moral law (*Stern*, 2013, p. 126-27). This section of the *Metaphysics of Morals* concerns feelings and so would be relevant only for creatures with sensibly affected feelings.³ Still, in such creatures, which includes humans, what we are seeing, again, is the correlative nature of right and value; in the absence of some value that is respected, the right cannot function as it should in prescribing duties.

Having a predisposition to self-respect may seem a long way from CI2's command to treat all rational beings with respect. For Kant, however, it is a journey of a single step. In the *Critique of Pure Reason*, he takes up the question of how it is possible to understand others as having minds. He is trying to defend his strategy of arguing from his analysis of how he thinks in the Transcendental Deduction to his critique of the Rational Psychologists' theory of thinking in the Paralogisms chapter. I quote the passage at length, because it is also key to his attempted demonstration of the freedom and so value of moral agents in *Groundwork* 3.

It must, however, seem strange at the very outset that the condition under which I think at all, and which is therefore merely a characteristic of myself as subject, is to be valid also for everything that thinks; and that upon a proposition that seems empirical we can presume to base an apodeictic and universal judgment, *viz*: that everything that thinks is of such a character as the pronouncement of self-consciousness asserts of me. The cause of this, however, lies in the fact that we must necessarily ascribe to things a priori all of the properties that make up the conditions under which alone we think them. *Now through no outer experience, but solely through self-consciousness, can I have the least representation of a thinking being. Hence objects of that sort are nothing more than the transfer of this consciousness of mine to other things, which thereby alone are represented as thinking beings.* (A 346/B 404-405, my italics)

³ I am grateful to several participants at a conference on Values at the University of Sheffield in March of 2023 for leading me to clarify this point.

In brief, since humans have only one window into the nature of thinking, namely their own thinking, they can represent another as a thinker only by using themselves as a model. Here is the one step needed to move from self-respect to respect for all moral agents:

1. The law of duty within forces a moral agent to have respect for her own being.
2. To think of others as having minds at all, an agent must use her mind as a model for them.
3. Since she uses her mind as a model and since she has the law of duty within, she represents others as having the law of duty within, which forces her to respect the beings that she thereby represents.

By combining his doctrine of the predisposition to self-esteem with his theory of representing other minds, we see that, for Kant, a moral agent not only has the law of duty within. She is also predisposed to respect all others whom she takes to be minded. Agents no more acquire respect for other moral agents from experience – which Kant thinks is just as likely to produce misanthropy (GMS, AA 04: 407, RGV, AA 06: 32) – than they acquire the law of duty from experience. The discussion of predispositions in the *Metaphysics of Morals* explains why having the law of duty and having respect for the being of others must be coeval in feeling creatures: Without respect for beings with the law within the putative law could not function as a law of duty.

4. How Does Kant Understand the Content of the Good?

I have based my arguments on an assumption about how Kant understands the content of the good, and I will now offer a retrospective defense of that assumption. In a sense Kant is clear about what the good is:

A rational nature exists as an end-in-itself. (GMS, AA 04: 429)

But we need to know which property of agents with rational natures provides the motivating ground for following C11. The question of what it is special about rational natures has been extensively studied and I begin with a standard reading offered by Henry Allison in his Commentary on the *Groundwork*. Allison considers three possible answers:

1. All minimally rational agents who have a capacity to set ends are ends in themselves.
2. Only agents with a good will are ends-in-themselves.
3. All rational agents with a capacity for morality are ends-in themselves. (Allison, 2011, p. 209)

Option 1 is familiar from the distinguished work of Christine Korsgaard (1986, p. 188) and Allen Wood (1999, p. 120-121). Still, as Allison notes, this well-known view can probably be set aside on textual grounds alone (2011, p. 216). Here are two passages that seem conclusive:

A person is a subject whose actions can be *imputed* to him. (MS, AA 06: 223)

This predisposition [to personality is] a special predisposition. From the fact that a being has reason it does not follow ... that this reason contains an ability to determine the power of choice unconditionally through the mere representation of the qualification of its maxims for universal legislation ... (RGV, AA 06: 26a).

The second citation is from the same note where Kant explains that the moral law must be “given within us.” Both indicate that moral personhood requires considerably more than the ability to set ends.

Both Allison (2011, p. 215) and Wood (1999, p. 120) dismiss the idea that good-willing agents are those deserving of respect, option 2, because the end in question must exist and Kant takes agents with a good will to be scarce and possibly non-existent (GMS, AA 04: 407). This seems too quick. A good will is celebrated in the *Groundwork*, from the opening sentence of section 1 to a discussion in *Groundwork 2* that follows Kant's summary of the relations among the three formulations of the CI. There, he notes that he is returning to where he began, with the good will, and explains that the existing end in question

can be nothing other than the subject of all possible ends itself because it is also the subject of a possible absolutely good will. (GMS, AA 04: 437)

The passage suggests that being an end-in-itself does not require an agent to have a good will, but only the potential for one. Although this option may seem to be a variant of Allison's preferred option 3, the moral capacity, it is not. On the fourth option, the underlying source of value is a good will; the value of the potential for a good will is derivative. By contrast, the moral capacity itself is valuable on option 3, and not just because it is a necessary condition for having the property that is valuable, *viz.*, an absolutely good will. I will offer reasons for thinking that Allison is right, that what makes rational nature an end-in-itself is the possession of the moral capacity, but I want to look first at some of Kant's language that makes the good will seem an apt candidate for the most fundamental moral value.

In discussing the value of the 'end-in-itself,' Kant's language ranges from flowery to florid. The end-in-itself does not have "relative worth", but "inner worth" (GMS, AA 04: 435); it is "infinitely above any price" (GMS, AA 04: 435); it has "absolute worth" (GMS, AA 04: 428), "dignity" (GMS, AA 04: 434), "sanctity" (GMS, AA 04: 435) and "sublimity" (GMS, AA 04: 439-40). Given these epithets, it can seem that the only thing capable of measuring up would be the one thing in the world or beyond it that is good without qualification, a good will. I argue below that the relation that Kant sees between the second and third formulations of the Categorical Imperative provides evidence that he takes rational nature to be an end-in-itself just by virtue of it possessing the moral capacity, but I think it is important to deal with Kant's colorful language. The absolute, incomparable, and sublime qualities of rational nature can give the appearance that Kant's ethics rests in some way on his teleology. Rational nature is an end-in-itself, because it is the end of nature. Good-willing agents are the *raison d'être* of the creation. Teleology is an important part of Kant's system and he argued that a realm of ends was nature's ultimate goal (KpV, AA 05: 548-49). Still, I will suggest in the last section that we can make sense of Kant's laudatory claims about rational nature within the context of his ethical theory and without relying on teleology.

The transition from the first to the second formulation of the Categorical Imperative focused on the basis or motivating ground for following the law of duty; the transition from CI2 to CI3 focuses on the necessary objectivity of morality. What I mean by the 'necessary objectivity' of morality is that it is part of the ordinary moral understanding that Kant is trying to make explicit that morality – if it is to exist or make sense – must be objective. Kant sounds the theme of the objectivity of morality again and again in the *Groundwork*. Just three pages into the Preface, he explains that according to the common idea of duty and morality, "a law, if it is to hold morally, i.e., as the ground of an obligation, must ... hold" not just for human beings but for all rational beings (GMS, AA 04: 389). What distinguishes moral obligations from mere preferences is that the former are objective; they bind agents regardless of their preferences. To explain the possibility of morality, Kant must explain how moral laws can be objective.

As he moves to CI3, Kant notes that he has not established the reality of morality, a task deferred to *Groundwork 3*, but simply assumed that it was possible. His arguments have been hypothetical: If there are objective moral laws, then they must be categorical (the argument that shows that any moral law must be universal and leads to the formulation of CI1); if there are objective moral laws, then their motivating ground must be an objective end, (the argument

that leads to the positing of rational nature as an end-in-itself and the formulation of CI2). He still cannot prove the reality of morality, but notes that

[o]ne thing, however, could have yet been done, namely: that the dissociation from all interest in willing from duty, as the specific mark distinguishing categorical from hypothetical imperatives be indicated in the imperative itself. (GMS, AA 04: 432)

And this leads to the third formulation:

The **principle** of every human will as a **will universally legislating through all its maxims** ... would be very well **fitted**, to be the categorical imperative, [because] **it is founded on no interest** and thus alone, among all possible imperatives, can be **unconditional**. (GMS, AA 04: 432)

The point of the third formulation – the formulation that Rawls made the foundation of his theory – is not to explain how moral laws bind or obligate agents. Kant is no contractarian. Agents are not bound by moral laws because they have agreed to them. They are bound by the law of duty within them and obligated to obey by their predisposition to respect rational nature in themselves and, through themselves, in others. What, then, is the problem for which CI3 is supposed to supply the answer?

CI1 commands agents to act only on maxims that they could will to be universal laws. Since moral laws are not laws of nature, they can become laws only through rational willing. Thus, if there are no principles of action that can be willed by agents, regardless of their subjective preferences, to be laws of action, then the set of objective moral laws will be empty and morality will be impossible. Conversely the possibility of morality requires that there be laws that agents can will independently of their particular interests. Since that is a requirement for something to be a moral law, an agent's action is morally permissible (can conform to law) only if she acts on a maxim that she could, as a legislator, make into a moral law.

To see why CI3 is relevant to the issue of the good will versus the moral capacity as the property that makes rational natures valuable, consider the relation between CI2 and CI3 on the two different interpretations. If the content of CI2 is that the moral capacity is an end-in-itself, then the conditions that are captured by CI2 and CI3 will each be a necessary condition for the other. Agents can have a moral capacity only if there is a community of agents under moral laws. Without a community, an agent has no way of understanding a principle as a moral law rather than as a preference. Moral agents require a community not just to engage in mutual aid, but to be able to make sense of morality, to be able to mark a distinction between preferences and obligations. From the other direction, it is impossible to have a community that legislates *moral* laws unless each member has the law of duty within and automatically respects beings that have the law within. Otherwise, you cannot have moral laws at all, but only dictates enforced by physical or social coercion and precepts based on self-interest.

Since acting with a good will requires the existence of objective moral laws, anyone with the potential for a good will must belong to a community of moral agents that collectively wills the laws. But the reverse relation does not hold. A community of agents under moral laws is possible just so long as each has the moral law within and automatically respects beings with the moral law within. It is not necessary that these capacities also make it possible for members of the community to have absolutely good wills, even though they do. Kant maintains that each of the formulae of the Categorical Imperative unites the other two within itself (GMS, AA 04: 436). What I have just argued is that CI3 unites CI2, as well as CI1, within itself if the property that makes rational agents ends-in-themselves is the moral capacity, but not if it is the potential for a good will.

We now have an explication of the property of beings with a rational nature that make them ends-in-themselves: It is the capacity for morality that includes having the law of duty within, a predisposition to respect beings with the law of duty within, and membership in a

realm of ends. To demonstrate the reality of morality, Kant must show that creatures with this capacity exist.

5. Demonstrations that Rational Moral Agents Exist

I will not defend Kant's demonstrations of the reality of morality. I present them for interpretive purposes, specifically, to illuminate the nature of his moral theory. Let us return to the passage to which Kant appends a note explaining that he is only postulating that rational nature exists as an end-in-itself, a proposition for which he will supply grounds in *Groundwork 3*. He asserts that each person necessarily conceives of his own existence as having intrinsic worth. What is to be established in last section is that

every other rational being also represents its existence in this way, *as a consequence of just the same rational ground that also holds for me.* (GMS, AA 04: 429, my italics)

As he presents the argument to come, it seems to involve a scope fallacy. Even if each person represents her existence as intrinsically valuable on the same ground that I regard my existence as intrinsically valuable, it hardly follows that each of us regards everyone *else's* existence as valuable on the same ground that she regards her existence as valuable. Since it would be surprising for Kant to commit such a fallacy, it is helpful to have a reading of the *Groundwork 3* argument on which he does not.

In *Groundwork 3*, Kant argues for an identity

A free will and a will under moral laws are one and the same. (GMS, AA 04: 447)

If his argument for the claim is sound, then he would be able to prove the reality of morality — that there are wills under moral laws — by proving that rational beings have free wills. And that is what sub-section 2 of the third section is supposed to show

Freedom must be presupposed as a property of the will of all rational beings.
(GMS, AA 04: 447)

Kant then offers his proof. Here are the first few moves:

1. One must prove it [freedom] as belonging universally to the activity of rational beings endowed with a will as such. 2. Now I say: every being that cannot act otherwise than **under the idea of freedom** is actually free, in a practical respect ... 3. Now I assert: that we must necessarily *lend* [*leihen*] to every rational being that has a will also the idea of freedom, under which alone it acts. For in such a being we conceive a reason that is practical, i.e., has causality with regard to its objects. 4. Now one cannot possibly think of a reason that would self-consciously receive guidance from any other quarter with regard to its judgments, since the subject would not then attribute the determination of judgment to his reason, but to an impulse. (GMS, AA 04: 448, my numbering and italics)

There are many confusing elements in this argument and much interpretive effort has been directed to the notion of acting under the idea of freedom, but I will focus on claims 3 and 4. Why would a subject who received external guidance for a “judgment” reject it as a judgment and label it as an “impulse” instead (claim 4)? Subjects have only one way of cognizing reasoning and that is from their own reasoning. In reasoning they are conscious in making a judgment and so having that thought, not conscious of thoughts intruding on reasoning from the outside. That is how they understand what reasoning is and why they cannot think of a reason that would self-consciously receive guidance from outside of itself. Hence, subjects are conscious of being free in that sense.

What about the odd locution, “lend,” in three? I take Kant literally. We lend other

rational beings the freedom from external interference of which we are conscious in our own case, when we project our mental activity onto them, as we must do to represent them as minded beings. On this reading, Kant would be providing just the argument for CI2 that he promised in *Groundwork 2*. Insofar as any individual takes herself to have intrinsic worth, because she recognizes herself to be free of external stimuli in rational action, she also takes all others whom she can understand as rational agents to have intrinsic worth on literally the same ground – *viz.*, her recognition of herself as a free agent. And, so for every rational agent.

Although it is widely agreed that Kant’s *Groundwork 3* argument fails to establish the reality of morality, there are many hypotheses about what goes wrong. Without wading into this interpretive thicket, I note two obvious problems. First, even if a subject is conscious of producing *judgments* independently of sensible impulses, that does not inform her that she can also produce judgments capable of guiding action independently of them. Second, even if humans can act independently of sensible impulses, that would not show that they can act on a principle with the content of the law of duty. One reason I highlight these objections is that the argument widely believed to replace the failed argument of *Groundwork 3* – the fact of reason passages of the *Critique of Practical Reason* – seems designed to meet them.

The argument of the *Second Critique* is surprisingly direct. As the Introduction explains:

[T]he first question is whether pure reason is sufficient by itself alone to determine the will or whether reason can be a determining basis of the will only as empirically conditioned. (*KpV*, AA 05: 15)

The question to be answered is whether reason, through its law of duty, can determine the will to action independently of any sensible desires. I follow Marcus Willaschek in seeing Kant’s demonstration as proceeding through a thought experiment (1991, p. 186). Kant invites the reader to consider two cases in order to prove to her – through her own deliberating – that her will can be moved by moral considerations. The crucial case is the second where a prince demands of someone:

on the threat of the ... penalty of death [by immediate hanging], that he give false testimony against an honest man whom the prince would like to ruin under specious pretenses ... He will perhaps not venture to assure us whether or not he would overcome that love [of life], but he must concede without hesitation that doing so would be possible for him. He judges, therefore, that he can do something *because he is conscious that he ought to do it*, and he cognizes [*erkennt*] freedom within himself – the freedom with which otherwise, without the moral law, he would have remained unacquainted. (*KpV*, AA 05: 30, my italics)

According to Kant, what happens when someone thinks herself into man’s position? Just prior to presenting the case, he claims that humans

become conscious directly [of the **moral law**] (as soon as we frame or draft or pose maxims of the will for ourselves) ... (*KpV*, AA 05: 29)

Kant assumes that in considering whether to bear false witness, his readers are conscious of the moral law. As he explains in the *Groundwork* when he claims that common human reason always has the moral law “before its eye,” he does not mean that the ordinary person is conscious of the abstract formula CI1, but that she is conscious of the concrete instance relevant to the case (*GMS*, AA 04: 402). Someone performing the thought experiment considers testifying falsely and thinks: “Unless I can will universal false testimony, I ought not to testify falsely.” From the recognition that she cannot will a world of false testimony, she infers: “I ought not to testify falsely.”

If Kant is right about how people think when putting themselves in the position the man whose prince demands false testimony – a huge if – then subjects participating in the experiment will be conscious of the moral law in two different ways. First, they are conscious of its content in the concrete case: I may testify falsely only if I can will a world of false witnessing.

Second, they are conscious of it as producing a change in their thinking, because they are conscious in inferring from this instance of the moral law to a judgment about what to do or, in this case, about what not to do: “I ought not to testify falsely.” Ordinary moral agents would not describe the change in their thinking in terms of principles of inference. A philosopher could, however, correctly describe the consciousness that Kant envisions as demonstrating that “act only on that maxim that your will could always hold at the same time as a principle in giving universal law” is an inferential principle that is used in practical deliberation. And this is the conclusion that he draws in §7, when he asserts that the moral law is a “basic law [*Grundgesetz*] of pure practical reason” (*KpV*, AA 05: 30).

With this result, Kant can finally answer the odd question from *Groundwork 2*. Humans necessarily judge actions through the law of duty, because that principle is a basic law of human practical deliberation. How he answers the question is by demonstrating to creatures who have the law of duty within that they do. The content of the law gives the determination or specification of the action: not testifying falsely. But the conclusion is not “I would rather not,” or “maybe I will not,” but “I ought not to testify falsely.” The “ought” indicates respect for the law of duty. Agents engaging in the thought experiments respect the law in their minds, even if they believe that their actions will probably fall short. And that is all that Kant needs. He is not trying to show something that is (obviously) false – that the moral “ought” always leads people to correct actions – but only that the human mind has the moral law within as a principle of deliberation and that it respects demands that derive from it.

As noted, Kant’s demonstration of the existence of morally capable agents, and so of the reality of morality, rests on granting him the outcome of the thought experiment that he expects. Again, my purpose in presenting the demonstration is not to defend it, but to fill out Kant’s theory of morality so that we can consider its status. The theory has two parts. First, arguments to the effect that the possibility of morality requires the existence of creatures who have the law of duty within, a predisposition to respect the beings of those who have the law within, and membership in a community of such beings. Second, an argument to the effect that there are morally capable beings, at least in so far that they have the law of duty within that they respect. My question for the last, brief section is “what kind of moral theory has Kant offered us?”

6. Kant’s Metaphysics of Moral Value

Kant refers to two kinds of goodness in *Groundwork 2*. Every imperative,

every practical law represents a possible action as good. (*GMS*, AA 04: 414)

Since there are two kinds of imperatives, there are two kinds of goodness. With hypothetical imperatives, the action is good, because it achieves a possible purpose, ultimately, happiness. But what good is the aim of categorical imperatives? Why should we believe that there is a second kind of goodness? On Rawls’s Kantian constructivism, moral value enters the world through a procedure. Because the procedure through which the basic structure is put in place is fair, actions in accord with laws sanctioned by the basic structure have moral value. They are good and good in a way that cannot be reduced to happiness.

For Kant, however, the source of moral value is different: existing beings with the moral capacity. Moral value is not real for him in the sense that it exists independently of rational agents and could be discovered by them. Moral value enters the world with the existence of creatures with the moral capacity. Creatures with the moral capacity could not discover that value, because recognizing and respecting that value are essential elements for having the capacity. Creatures with the moral capacity take that capacity – rational nature – to have intrinsic worth, to have

dignity and to be beyond any price. In using such language, Kant is not, I believe, appealing to teleology, but continuing his project of explicating ordinary moral understanding. It is part of ordinary moral understanding that morally capable creatures have a special value or status. That value is what confers goodness on actions that respect it and evil on actions that do not. Here we need a delicate balance. One element of the moral capacity is the ability to recognize moral agents and respect their value, but that does not mean that moral value exists only by virtue of being recognized and respected⁴ or that it exists only from the perspective of moral agents.

To see the difference between Kant's position and a perspectival theory, recall Mill's unfortunate discussion of higher and lower pleasures. Mill wants to offer an absolute judgment: It is better to be a human being unsatisfied than a pig satisfied. Mill defends his claim by arguing that only beings who are capable of higher pleasures can rank the pleasures and so pronounce on the sorry state of a satisfied pig (Mill, 1861/1991, p. 139-40). Unlike Mill, Kant does not argue that the human perspective should have dominion over a pig perspective – at least when he is not doing teleology – but he is not a relativist either. His view is not that, from the perspective of humans, moral value exists, but from the perspective of pigs, it does not. Kant's position is that it is part of ordinary moral understanding that creatures with the moral capacity have value, and that since those creatures exist, moral value – as we understand it – exists. The world contains a second kind of goodness, moral goodness, because the world contains creatures with the moral capacity. Even though pigs lack the ability to recognize the moral capacity and its value, they inhabit a world where moral good and evil exist.

Bibliographic References

ALLISON, Henry E. *Kant's Groundwork of the Metaphysics of Morals: A Commentary*. Oxford: Oxford University Press, 2011.

GREGOR, Mary, trans. and ed.. *Immanuel Kant. Practical Philosophy. The Cambridge Edition of the Works of Immanuel Kant*, P. Guyer and A. W. Wood, (gen. eds.). New York: Cambridge University Press, 1996.

KANT, Immanuel. *Kants gesammelte Schriften, Akademie Ausgabe, edited by the Koeniglichen Preussischen Akademie der Wissenschaften*, 29 vols. Berlin: Walter de Gruyter and predecessors, 1900-.

KORSGAARD, Christine. Kant's Formula of Humanity. *Kant-Studien*, v. 77, n. 2, p. 183-202, apr. 1986.

MILL, John Stuart. On Liberty and Other Essays. In: GRAY, John (ed.). *Oxford World Classics*. Oxford: Oxford, 1861/1991.

PLUHAR, Werner, trans. *Immanuel Kant. Critique of Pure Reason: Unified Edition*. Indianapolis: Hackett, 1996.

PLUHAR, Werner, trans. *Immanuel Kant. Critique of Practical Reason*. Indianapolis: Hackett, 2002.

PLUHAR, Werner, trans. *Immanuel Kant. Religion within the Bounds of Bare Reason*. Indianapolis: Hackett, 2009.

⁴ I am grateful to Robert Stern for raising questions about an earlier draft that led me to clarify this issue.

RAWLS, John. Kantian Constructivism in Moral Theory. In: FREEMAN, Samuel (ed.). *Collected Papers*. Harvard: Harvard University Press, 1980/1999.

RAWLS, John. *Lectures on the History of Moral Philosophy*. Barbara Herman, (ed.). Harvard: Harvard University Press, 2000.

STERN, Robert. "Kant, Moral Obligation, and the Holy Will." In *Kant on Practical Justification*, Mark Timmons and Sorin Baiasu, (eds.). Oxford: Oxford University Press, p. 125-152, 2013.

WILLASCHEK, Marcus. *Praktische Vernunft: Handlungstheorie und Moralbegründung bei Kant*. Stuttgart: J. B. Metzler, 1992.

WOOD, Allen W. *Kant's Ethical Thought*. Cambridge: Cambridge University Press, 1999.