

1 Introduction

1.1 Why Study Kantian Ethics?

Readers and students of Kant's ethics sometimes face a frustrating paradox. On the one hand, Kant's moral theory seems to describe a moral principle that is, on its surface, almost intuitive. Familiar themes of reciprocity, impartiality, and fairness play a central role, and one is left, perhaps, with the sense that the moral code described is, at the end of the day, something not too far removed from the golden rule or some similar preschool edict. On the other hand, as soon as one begins to study the theory carefully, one is beset with difficulty. Kant's argumentation is technical and dense, and it takes a puzzling path through intricate analyses of intentional action and moral judgment just to arrive at a statement of the moral law that is itself difficult to unpack and apply in every case – a moral law that apparently contains several different versions or formulations within itself. All of this is to say nothing of the exactitude and strenuousness with which we are apparently supposed to live our moral lives and the unwavering devotion one ought to have to morality itself. After all of that, a person could be forgiven for wondering if she might not be better off just sticking with the moral code she learned in preschool, after all.

The Kantian reply is straightforward: the moral code you learned from your preschool teachers or parents, though it has perhaps served you well, came from an external authority, not from you. True, there might be a sense in which you sometimes endorse it or parts of it, but you were not its legislator, and this makes all the difference. What is more, there is (if we're honest) a puzzle about *why* you follow it at all, if and when you do. Maybe you have simply internalized the authority that your teachers and parents had over you so long ago. Or maybe you discovered along the way that following these general rules was the best way to avoid trouble, get by, or bring about your own well-being or that of those you care most about. But then what about those times when – as so often happens – it doesn't actually succeed in accomplishing these ends? Many a former preschool student has wondered, in later years, what the point of following all these rules is when the ones who flaunt them or treat them as mere fiction seem to get so much further ahead.

What Kant's moral theory has to offer is a way to make sense of our moral obligations from the ground up, as it were. Kant offers an account of morality whose validity does not depend on external authority, or on the positive consequences it might bring to oneself or others. It is a morality that each of us authors herself, in light of certain rational requirements and in light of the fact that we share a moral community with other agents. The aim of this text is to provide an overview of this moral principle, including its foundations and application.

1.2 Overview of the Text

The text begins with a survey of some of the major themes that run throughout Kant's moral philosophy. These, I hope, will provide some guidance and context as the text moves on to a study of Kant's arguments concerning the foundations of morality. Although the text serves as an introduction to the arguments that Kant provides regarding the moral law, it is not a commentary on the text that many students read when first studying Kant's moral philosophy, the *Groundwork for the Metaphysics of Morals (Groundwork)*.¹ Nevertheless, the section headings in this text should provide some guidance with regard to where overlap does occur.

The text continues with a discussion of Kant's account of moral and nonmoral willing (Section 3), which is central to his overall argument. It moves from there to a discussion of the first two formulations of the categorical imperative: the formula of universal law and the formula of humanity (Sections 4 and 5, respectively).² The text then discusses Kant's account of moral motivation (Section 6), which leads naturally to a discussion of autonomy and freedom of the will (Sections 7 and 8, respectively). In Section 9, I attempt to answer a few of the more common objections to Kantian ethics.

I have tried, as far as possible, to avoid jargon in an effort to make the text accessible to a wide range of readers. Nevertheless, Kant's argumentation can often be difficult, no matter how straightforwardly one tries to present it, and one Kantian argument can often admit several plausible interpretations. Alas, a text like this cannot do justice to the various interpretations and arguments that Kant scholars have offered over the years, though I have attempted to highlight moments in the discussion where there is considerable debate among Kant scholars. Footnotes throughout the text also provide guidance for readers interested in pursuing a topic in more detail.

2 Themes in Kantian Ethics

2.1 Universality and Necessity

Universality and necessity are central concepts throughout Kant's philosophical system. In his moral philosophy, they play a crucial role as part of his argument

¹ Commentaries on *Groundwork* certainly exist (see Allison, 2011; Guyer, 2007; Timmermann, 2007). Uleman (2010) and Wood (1999) offer useful commentaries on Kant's ethical thought as a whole.

² Among Kant scholars, there is some variation in how the various formulations of the categorical imperative are counted and categorized. For example, Paton (1971) takes the formula of the law of nature to be a version of the formula of universal law and the formulations of autonomy and the realm of ends as falling under the formula of humanity. Timmermann (2007) presents the formula of universal law as the overarching formulation, of which the others are subspecies.

about the nature of moral laws and moral obligation. Among Kant's foundational premises is a definitional point – that a moral law, if it is truly a *law*, must hold universally and necessarily (*GMS*, 4:389). Kant takes this to be uncontroversial, and his argumentative strategy will be to investigate what kind of moral code could possibly have these features of universality and necessity. Ultimately, he will conclude that only the principle of morality that he calls the categorical imperative can fit this description. Relatedly, our obligation to abide by the moral law described by the categorical imperative is also universal and necessary. In other words, the categorical imperative describes the content of a moral law that all rational agents are bound by necessarily. The moral law is thus necessary and universal both in the content that it describes and in its account of agents' moral obligation. A large part of Kant's argument – and the discussion in this text – will be devoted to arguing for these claims.

Perhaps because of Kant's insistence upon universality and necessity with respect to the obligation it imposes, he is sometimes cast as a kind of unwavering rule fetishist, most infamously in examples like those concerning the “murderer at the door,” in which an agent is faced with a choice of whether to unwaveringly follow a rule (e.g. against lying) or break it in order to save a friend's life or for the sake of some other worthy goal. It would be an exaggeration to say that this impression is entirely unfounded, and there is no simple or straightforward way to respond to these objections and counterexamples.³ Still, the objection gets a good deal of extra steam from the assumption that Kant is offering a fixed system of preexisting commandments or rules. However, Kant's moral system is not, despite its insistence upon necessity, a system of *rules*. Instead, Kant sets out a test of *permissibility*: The categorical imperative tests whether the principle or plan of action that one has set out for oneself is permissible or impermissible. Now, of course, certain types of principles – those that involve the subjugation of others, for example – will turn out to be impermissible whatever the circumstance. Still, it is central to Kant's account that the agent considers her own proposed principle of action and whether this would be permissible. Kant's moral system, though at times rigorous and demanding, does not describe a moral world in which obligation is imposed upon agents as a set of commandments from a stern and unforgiving external authority. Instead, Kant thinks that morality can only be imposed on an agent from within. It is and must be self-given. This is Kant's notion of autonomy.

³ For a detailed discussion on the question of the murderer at the door, especially regarding how to understand the notion of a lie, see Mahon, 2003, 2009.

2.2 Autonomy

Often, we use the term “autonomy” to describe self-sufficiency or self-determination, as when we discuss patient autonomy in healthcare settings or the importance of developing children’s sense of autonomy in the context of early education. For Kant, the term has a much narrower meaning: it refers to an agent’s capacity to simultaneously be subject to and the legislator of the moral law. On the Kantian view, in other words, it is crucial that agents are the source of the moral prescription that they themselves follow. This does not mean, however, that agents can arbitrarily decide upon the content of the law. As Kant puts it, the agent is the author of the “obligation in accordance with the law” but not of the law itself (*MS*, 6:227).

This view sets Kant’s theory apart from theories according to which a moral code is given and enforced by an external authority, whether this is a divine authority or some other, worldly authority. But it also sets Kant apart from philosophers who argue that moral knowledge and moral motivation stem from an emotive response or moral sense.⁴ Although that sort of moral sense would be “internal” to the agent, Kant argues that this does not put the agent in charge any more than she is in charge of any other impulse or feeling – say, of her hunger or her sadness. This brings us to a key point about autonomy – in order to be the author of the law, an agent’s *reason*, not her feelings or her desires, must be legislative.⁵

Importantly, however, autonomy does not mean that the moral law is arbitrary or “up to us” in the sense that each individual can simply decide the difference between right and wrong on the basis of what is convenient or brings her the most pleasure, for example. That would be to make morality a matter of instrumental reason, that is, of reasoning about which actions best promote one’s interest. Reason plays an important role in Kant’s ethics, but it is not *instrumental* reason that plays this role. Crucial to his account of autonomy – and morality in general – is Kant’s view that moral agents are capable of another type of reason altogether – a type of reason that is able to consider whether our actions are moral, without thinking about whether they are in our own interest. Now, of course, these two types of reason often find themselves at odds with one another; the most prudent course of action from a self-interested perspective is not always (alas!) the right action, considered morally. This, too, will be a theme in Kant’s ethics: we are all too aware of the fact that doing the right thing can often involve a degree of pain or sacrifice. But the very fact that we recognize

⁴ Kant has in mind moral sense theorists, for example, Shaftesbury, Hutcheson, and Hume.

⁵ Reason is legislative in the sense that it issues the moral principle to be followed. Moral motivation, for Kant, is a trickier subject. See Section 6 for further discussion.

this tension, Kant would say, gives us a clue that we are capable of recognizing constraints on our actions whose source is something other than self-interest.

2.3 Egalitarianism and Impartiality

The discussion that follows will often emphasize Kant's egalitarian commitments. At the root of these is Kant's conviction that every rational agent is an end in herself and should never be treated as a mere means to an end (see Section 5). This gives all rational agents a kind of dignity that can never be exchanged or traded for the sake of other ends (*GMS*, 4:435). This commitment to the fundamental, inalienable dignity of all rational agents means that Kant's egalitarian commitments can often be stated as prohibitions against making an exception for oneself or using others merely as a means toward one's own ends. Kant is thus in some sense the forerunner of contemporary theories that emphasize treating others in ways that they could rationally consent to from an impartial standpoint or treating others in ways that they could not reasonably reject.⁶ Crucially absent from Kant's egalitarianism is the approach taken by many consequentialist theories, according to which each relevant individual has one "vote," as it were, when it comes to deciding which available course of action will bring about the most net utility. Indeed, this approach easily violates the prohibitions described above, since on this account, one person's dignity may sometimes need to be sacrificed for the sake of the greater good.⁷

2.4 Freedom and Nature

Finally, it is important to note that on the Kantian view, we are both moral beings and sensible beings. This means that we are capable of understanding and giving ourselves the moral law *and* that we are simultaneously subject to feelings of pleasure and displeasure and the desires and inclinations that develop from these feelings. We occupy, as it were, two different worlds and find ourselves continuously navigating between them. This has several important implications for Kant's ethics as a whole. First, it will mean that we will experience morality as a kind of obligation or constraint. Though self-imposed, morality will nevertheless tell us that some of the things we plan or hope to do on the basis of inclination are impermissible. This brings us to a second point: moral life, for sensible creatures like us, will involve a host of different feelings; when we discover that a proposed plan of action is impermissible, we experience a kind of "pain" or even "humiliation" (*KpV*, 5:74). This, in turn, inspires a feeling of awe or reverence for the moral law (*KpV*, 5:74). Thus, although

⁶ See, for example, Rawls, 1980; Scanlon, 1998.

⁷ See especially Rawls, 1971, p. 24, (rev. ed.) on this point.

feeling cannot serve as the *foundation* for a universal and necessary principle of morality, Kant's account of the moral existence of sensible yet rational creatures includes a good deal of feeling indeed.

We are sensible beings, and we are also finite beings. We have desires and ends we cannot always accomplish on our own, and we often need the assistance of others, whether this is because we want to learn a skill or because we suffer and need help. The fact that we are sensible and finite beings will thus also serve as an important premise in Kant's arguments regarding our duties of assistance toward others. Taken together, the preceding observations indicate that our sensible nature can often be the source of need or frustration. Although Kant sometimes suggests that the best or most rational thing would be to wish to be free from inclination altogether (*GMS*, 4:454), he does not, in my view, mean to renounce our sensible nature altogether. In the first place, it would be an idle wish. Furthermore, however, it is because we are sensible that we can take disinterested pleasure in beauty and experience happiness when our interests are fulfilled. Kant is not a Stoic: he thinks that our general well-being – which he would term “happiness,” though the Stoics would not – is important to us; it is a conditional good (i.e. it is good as long as it is consistent with morality). Rather than seeking to renounce our sensible nature, then, Kant is best understood as seeking a way to make happiness consistent with virtue, as far as this is possible.⁸

3 The Moral Law and the Will

Having considered some of the major themes in Kant's ethics, we are now in a position to examine Kant's foundational arguments regarding the moral law and moral willing. Kant sets out by posing a question: If there is such a thing as a universally valid moral law, what would be the characteristics of that law? A further question follows immediately: Given the characteristics of a universally valid moral law, what characteristics must moral *agents* have? Specifically, how should we describe the will of the agent who is bound by this law? For sake of exposition, we will discuss these questions independently of one another, but there is an important sense in which they are intimately connected in Kant's arguments regarding morality. In particular, Kant's arguments regarding the content of morality are only provisional until he can also demonstrate that moral agents are actually capable of and subject to the moral law that he describes.⁹

⁸ Kant seeks to accomplish this, in large part, via his account of the so-called “highest (derived) good,” discussed briefly in Section 9 of this text.

⁹ This is the structure of the argument in the *Groundwork of the Metaphysics of Morals*. See Section 8 for more detail about the argument and how it shifts over time.

Kant takes it as a foundational starting premise that a moral law must hold with unconditional necessity. This is a general, definitional point about the characteristics of a *law*: If something is to count as a *law* – and not merely as a guideline, for instance – it must be the kind of principle that applies universally (*GMS*, 4:389). Kant thus takes the necessary character of the moral law to be contained within the notion of lawfulness. Hence, insofar as we are looking for a description of the moral *law*, then, we are looking for an account of morality that is able to account for this universality. As we will see in what follows, Kant thinks that this rules out many of the more popular contending accounts of morality.

The moral law – if indeed there is such a thing – applies universally and necessarily. But to *whom* does it apply necessarily? We shouldn't assume, for example, that because it is universal and necessary, the moral law applies to all living creatures, all sentient creatures, or even all creatures with only limited rational capacities. While we might think that nonhuman animals deserve some type of moral consideration, we tend not to think of our pets, for example, as fellow moral agents capable of recognizing and following moral guidelines.¹⁰ This is why many, if not most, of us find it perfectly acceptable to manage and condition our pets' behavior by putting them in kennels or on leashes and not with appeals to their sense of fairness or reciprocity. Insofar as there is a universally valid moral law, Kant will argue that the beings it applies to will have to have the capacity to recognize laws or principles and to let their action be guided by those laws or principles. In Kant's terminology, the moral law applies to agents who are able to "act according to the representation of a law" (*GMS*, 4:412). Any creature that is capable of acting according to the representation of a law is, in theory, capable of morality and subject to its requirements. Kant's notion of acting under the representation of a law is central to his account of action in general – that is, to his account of both moral action and nonmoral action. It will thus be helpful to examine this capacity in more detail.

Any instance of intentional action involves, in the Kantian terminology, an *objective determining ground* and a *subjective determining ground*. The objective determining ground is the guideline, principle, or law that the will follows, depending on its aims and circumstances (e.g. *GMS*, 4:400). The subjective determining ground provides the motivation for the will (e.g. *GMS*, 4:413n).¹¹

¹⁰ For further discussion of Kant's arguments regarding our duties concerning nonhuman animals, see Kain, 2010; Timmermann, 2005; and Varden, 2020. A novel Kantian account is offered in Korsgaard, 2018.

¹¹ The subjective determining ground is thus not the same as a "maxim," which Kant describes as a "subjective principle of willing" (*GMS*, 4:400n). Rather, it is what moves an agent to adopt a maxim.

To get a clearer grasp on how objective and subjective determinations of the will work in concert, it may be useful to begin with an example of nonmoral willing. Imagine a dressmaker who wants to make a strapless dress. She knows that in order to do this, she will need to add structure to the bodice of the dress, say, by inserting boning. Here, the objective principle informing the dressmaker's choice of action is something like: "In order to make a strapless dress, you need to add boning to the bodice." The objective principle is objective because it applies to any dressmaker who aims to make a strapless dress. Still, although the principle is objective, it is not authoritative or action-guiding for every agent, for the simple reason that only a very small subset of agents is, at any given moment, concerned with making a strapless dress. In order for this particular objective principle to apply to an agent, she must have the relevant aim or end. The desire to make a strapless dress is the subjective determining ground of the agent's will. Now, we can see how objective and subjective determining grounds work together in the case of nonmoral willing: the objective determining ground provides principles that any agent with a particular subjective determining ground should follow in order to accomplish her ends.

Notice that in the nonmoral case, neither the objective nor the subjective determining ground of the will is universally and necessarily valid. This is probably easier to see in the case of the subjective determining ground: the principles that describe dressmaking will do nothing to move the agent hard at work building a chest of drawers or filing her taxes. But even the objective principles that apply to agents who *do* adopt certain ends are changeable: facts about the agent and the world can affect the objective principle that guides an agent's choice of action. Before the advent of various textile innovations like elastic, boning might have been essential to the construction of a strapless dress. Today, such innovations provide other options, and this changes the objective principle a person must follow when she sets out to design such a garment.

It is worth pausing here to note that this contingency regarding objective principles is not just a feature of textile design or other "imperatives of skill," as Kant will call them (*GMS*, 4:415). It is also a feature of more general aims and projects, most notably our pursuit of happiness. Many of our actions aim at happiness, but it is impossible to pin down an objective principle that will invariably lead to the attainment of happiness. This is because agents and the world they live in are unpredictable and constantly changing. A principle of saving money for later in life is perhaps a good general objective principle of happiness, but not if one's life is cut short unexpectedly. In that case, it might have been better from the point of view of happiness to spend one's money while one had the opportunity. And it is not at all uncommon to find that some of the things that we think will bring us happiness ultimately fall short in this

regard, sometimes because we ourselves change over time. These observations about the variability and unpredictability of any objective principle regarding happiness lead Kant to reject the idea that the pursuit of happiness – that of our own or that of general happiness – could be the foundation of a moral *law* (*GMS*, 4:418). Kant acknowledges that we do, in fact, all pursue happiness as an end, but the principles guiding its attainment are themselves far from necessary or universal. Of course, even if the attainment of happiness were predictable and determinate, it still could not serve as a foundation for the moral law, since each of us only seeks our *own* happiness. The harmony and universality required by a moral *law* would be thrown into chaos by a principle founded on *general* happiness (*KpV*, 5:28).

In Kant's view, *any* kind of intentional action requires subjective and objective determining grounds. So far, our examples have only concerned nonmoral action, which operates according to subjective and objective determinations of the will that are contingent upon facts about the world or upon the aims and desires of particular agents. But *moral* willing is also a kind of willing, and it must have its own objective and subjective determining grounds. As we have already noted, Kant thinks that a moral law applies necessarily. So we can think of Kant's project – particularly in the *Groundwork of the Metaphysics of Morals* and the *Critique of Practical Reason* – as a search for and explication of the objective and subjective determining grounds of the will when it engages in *moral* willing. The aim, in other words, is to find objective and subjective determining grounds of the will that could carry with them absolute necessity and universality.

But here we run into a serious challenge. We can readily locate determining grounds of the will for nonmoral willing precisely because facts about the world and agents' aims make these clear to us. Facts about gravity, the shape of the human body, and the physical properties of fabric combine to generate a rule about how to construct a strapless dress. Note, however, that it is precisely these facts about the world that make the principle contingent. Similarly, a person's desire to make a strapless dress explains immediately and straightforwardly why she is following the principle to insert boning into a dress. But, again, the desire to make a strapless dress is a contingent matter – indeed, partially contingent on what is fashionable in any given context! Thus, the very same features of the world and agents' desires that make nonmoral action relatively easy to explain also *exclude* them from consideration as the basis of necessary and universal *moral* willing. In order to find objective and subjective determinations of the will for moral willing, we will thus have to look elsewhere.

4 The Moral Law: The Formula of Universal Law

We are now in a position to consider Kant's explication of the moral law – the principle, or objective determining ground of the will when the will is guided by morality, as opposed to inclination and desire. Above, we saw that in the nonmoral case, agents' aims, combined with empirical facts about the world, generate the rules that agents use to guide their actions. But morality must apply universally and necessarily, and agents' aims and empirical facts about the world are wholly contingent matters.¹² So we seem to be at an impasse, since everything that typically provides content to the rules or laws that a will acts upon is excluded by virtue of its contingency. However, Kant argues that one thing still remains even after we have abstracted any material aim: something we might call the form of lawfulness as such. As he puts it, "Since I have robbed the will of all impulses that could arise for it from following some particular law, nothing remains but as such the universal conformity of actions with law, which alone is to serve the will as its principle" (*GMS*, 4:402).

Stated in these terms, Kant's thinking is perhaps a bit murky. It may be helpful to think of Kant as saying something like the following: "Since everything that normally ties rules or laws to the will is excluded as contingent in the case of morality, the only thing left is to see whether the very notion of a necessary and universal law might *itself* give content to a moral law."¹³ And Kant thinks it does: the principle that emerges, he argues, is the following: "I ought never to proceed except in such a way *that I could also will that my maxim should become a universal law*" (*GMS*, 4:402, emphasis in original). This is Kant's famous categorical imperative. It is an *imperative* in the sense that it is an action-guiding, objective determining ground of the will, and it is *categorical* because it applies necessarily and universally to any agent capable of acting under the representation of a law or principle, no matter their particular aims or interests.

It is worth keeping a few things in mind at this stage of Kant's argument. First, at this stage in the *Groundwork*, Kant is not yet arguing that this moral law "exists" or that it actually applies to us. That will require more argument (see Section 8). Here, Kant's only assertion is that *if there is a moral law*, it would have to have these characteristics. Imagine, by way of outlandish example,

¹² Terminologically, Kant thus often distinguishes between a *rule*, which can admit of variation and contingency, and a *law*, which cannot. However, he does not always abide by this distinction.

¹³ It bears pointing out that some scholars have worried about a gap in Kant's argument at this stage. See, for example, Allison, 1996 and Aune, 1980. Korsgaard (1996, pp. 61–64) offers a reconstruction of Kant's argument. Timmermann (2007) argues that the gap can be closed. See also Gaut and Kerstein, 1999.