

## Kant and Consequentialism in Context: The Second *Critique's* Response to Pistorius

Michael Walschots

\*Penultimate version. Final version published in:

*Archiv für Geschichte der Philosophie*. 103 (2), 2021. pg. 313–340.

<https://doi.org/10.1515/agph-2017-0132>

### Introduction

It is fairly standard practice to contrast Kant's moral philosophy with consequentialist or utilitarian approaches to ethics. This is because Kant is thought to have offered a kind of moral theory that is, at its core, fundamentally distinct from consequentialism.<sup>1</sup> At the same time, over the years a number of commentators have suggested that Kant's moral theory is compatible with consequentialism to varying degrees. For example, more than 20 years ago Richard Hare asked: "Could Kant have been a Utilitarian?" and he concluded that "Kant [...] *could* have been a utilitarian, though he *was* not. His formal theory can certainly be interpreted in a way that allows him – perhaps even requires him – to be one kind of utilitarian" (Hare 1997, 148), namely what Hare calls a "rational-will utilitarian" (see *ibid.*, 151). David Gmmiskey took an even more drastic approach and attempted to develop a kind of *Kantian Consequentialism* (see 1996), and more recently Derek Parfit sought to combine Kantian and utilitarian ethics into his so-called "Triple Theory" (see Parfit 2011). These attempts to make Kantian ethics and consequentialism compatible in turn encouraged others to argue that they are fundamentally *incompatible* ethical theories.<sup>2</sup>

What all the discussions surrounding Kant and consequentialism have in common, however, is that they explicitly take place from the perspective of systematic moral theory; the

---

<sup>1</sup> Kant's moral theory is often categorized as a "deontological" view, but Jens Timmermann has recently argued that the word "deontology" is relatively meaningless or at least unhelpful. Timmermann persuasively suggests that we abandon using the word deontology, that we stop classifying Kant as a deontologist, and also that we discard the distinction between deontological and consequentialist theories of ethics (see Timmermann 2015).

<sup>2</sup> See e.g. Rohs (1995) and Timmermann (2005).

question is never whether Kant *was* a utilitarian or presented a consequentialist moral theory (indeed, in the passage above Hare explicitly states that Kant *was* no utilitarian), rather it is a question of the extent to which Kant's moral theory is compatible with consequentialism on a systematic level. This is of course for good reason. First, it is rightly assumed that Kant himself in no way attempted to put forward a consequentialist moral theory. Second, Kant was indeed "oblivious" (see Timmermann 2014, 240) to the first and earliest attempt to work out a utilitarian ethical theory, namely that contained in Jeremy Bentham's *An Introduction to the Principles of Morals and Legislation* (1780), even though it was produced around the same time as Kant's first major works in ethics, i.e. the *Groundwork for the Metaphysics of Morals* (1785) and the *Critique of Practical Reason* (1787).<sup>3</sup> At the same time, this understandable distance from approaching the question of the compatibility of Kant's ethics and consequentialism from a historical point of view has had an unfortunate consequence: it has resulted in interpreters neglecting to ask whether Kant himself had reasons for why he considered his own view to be distinct from a broadly consequentialist moral theory.<sup>4</sup> As I hope to show in this paper, there is

---

<sup>3</sup> This is not to ignore the fact that there are even earlier ethical theories that have features we might now consider to be consequentialist, and of which Kant would have been aware. The most obvious candidate is Epicureanism, but as Joachim Hruschka has argued there is also reason to believe that Leibniz was one of the, if not *the*, first to formulate the greatest happiness principle (see Hruschka 1991). Francis Hutcheson is another figure who plays an important role in the history of consequentialism, and whose works Kant read in translation. It is thus likely that Kant was familiar with at least these early, proto-consequentialist views. As I explain below, my aim in this paper is not to discuss all the early forms of (proto-) consequentialism, with which Kant may have been familiar. This is a worthwhile task, but is one that would require a separate paper. For a helpful overview of utilitarian thought before Bentham, see Heydt (2014).

<sup>4</sup> It should be noted at the outset that I am not dealing with a specific version of consequentialism throughout the discussion of this paper, but only a very general consequentialist view of the kind defined by Samuel Scheffler: "Consequentialism in its purest and simplest form is a moral doctrine which says that the right act in any given situation is the one that will produce the best overall outcome." (1988, 1) The important, core feature of such

reason to believe that Kant did indeed have such reasons. I will argue that Kant offered reasons for why his moral theory is distinct from consequentialism by way of responding to an important criticism of his moral theory, namely the allegation that his principle of moral appraisal is implicitly consequentialist. This allegation was most famously made by figures such as Arthur Schopenhauer and J.S. Mill and has been repeated by numerous other commentators since. Even more importantly for my purposes is that this allegation was also put forward during Kant's lifetime in two separate instances: in an early review of the *Groundwork* by H.A. Pistorius as well as in a book length discussion of Kant's *Groundwork* by G.A. Tittel entitled *Kant's Moralreform*. In the following I demonstrate that Kant knew of both Pistorius's and Tittel's works and suggest that an indirect response to the allegation of consequentialism is contained in the *Critique of Practical Reason*. It is Kant's response to this allegation, I will argue, that illustrates he had views of his own regarding why his moral theory was distinct from consequentialism.

The following discussion is divided into three sections. In section (I), I provide an account of the allegation of Kant's implicit consequentialism. This allegation has two main versions: an egoist version and a utilitarian one. I explain the features of each version before outlining the two instances Kant responds to in the second *Critique*, namely those put forward by Pistorius and Tittel. In section (II), I argue that Kant indirectly responds to the allegation by distinguishing his own conception of the good and moral appraisal from a view he calls "empiricism of practical reason," a broad category of moral theory that he considers consequentialism to fall under. In section (III), I return to the question of what Kant took to be the core difference between his moral theory and a broadly consequentialist conception of ethics. Although the second *Critique* confirms what others have argued, namely that what distinguishes these approaches to ethics is their conception of value and their moral principle, I argue that Kant himself saw the *fundamental* difference to be one that concerned methodology and the priority of the right to the good.

My overall aim in this paper is thus to identify what Kant takes to be the primary difference between his own moral theory and one that can be classified as broadly consequentialist in spirit. I accomplish this by giving an account of an important, early criticism

---

a view for my purposes is that moral appraisal judges actions, maxims, etc. as morally good if and *because* they produce the best overall outcome, no matter how this outcome is construed.

of Kant's moral theory and his subsequent response to it in the second *Critique*. I do not attempt to reconstruct all of Kant's arguments against consequentialism, which are contained in many other texts, especially but not limited to the *Groundwork*. I also do not attempt to give an account of all the early (proto-) consequentialist moral theories with which Kant might have been familiar, and which he may have in mind when he discusses such views in other places. A complete discussion of Kant and consequentialism from a historical point of view would require evaluating a much broader context and taking many more of Kant's texts into consideration. I must therefore defer these tasks to another occasion.

### I. The Allegation: Kant's Implicit Consequentialism

Ever since Kant first published his views on moral philosophy his critics have found a number of its features difficult to understand. Perhaps one way of ascertaining which features are the most perplexing is to look at which objections and reactions find themselves repeated by various authors over the course of the last 200 years of scholarship and criticism. One such reaction, and one which has been discussed surprisingly little in the secondary literature despite the frequency of its appearance, concerns Kant's conception of moral appraisal<sup>5</sup>, i.e. his account of the way in which we judge the morality of actions, maxims, and characters<sup>6</sup>, and the principle we employ when doing so. According to some commentators, Kant's conception of moral appraisal is implicitly consequentialist. In this section I explain two different versions of this allegation: what I call the "utilitarian" and the "egoist" versions.<sup>7</sup>

---

<sup>5</sup> The German term for "appraisal" here is *Beurteilung*, which has been variously translated as "estimation," "appraisal," and "judging." I adopt "appraisal" to remain consistent with current leading translations of the *Groundwork* (Gregor and Timmermann) and the second *Critique* (Gregor).

<sup>6</sup> As I discuss further below, Kant identifies these items (actions, maxims, and characters) as the primary objects that can be judged as morally good or evil.

<sup>7</sup> Although different in flavour, these versions of the allegation are of the same general type in that they both attribute to Kant the view that moral judgement takes place by means of determining the *consequences* that the universalization of a maxim might have on one's own (the egoist version) or collective (the utilitarian version) well-being. Although it is the utilitarian

John Stuart Mill put forward what is likely the most well-known version of this allegation, at least in the English-speaking world. In *Utilitarianism*, Mill claims that Kant “fails, almost grotesquely, to show that there would be any contradiction, any logical (not to say physical) impossibility, in the adoption by all rational beings of the most outrageously immoral rules of conduct” (Mill 1969, 207). According to Mill, all Kant does show about certain rules of conduct “is that the *consequences* of their universal adoption would be such as no one would choose to incur” (ibid.). Mill summarizes his view later in the text: “To give any meaning to Kant’s principle, the sense put upon it must be, that we ought to shape our conduct by a rule which all rational beings might adopt *with benefit to their collective interest*” (Mill 1969, 249).

At first glance, this criticism might come across as misguided given Kant explicitly distances himself from such a position during the first presentation of his theory of moral appraisal in the *Groundwork*. Kant argues there that all one needs to do in order to determine whether an action is in conformity with duty is ask oneself: “would I indeed be content that my maxim [...] should hold as a universal law.” (4:403)<sup>8</sup> When performing this universalization test, as it has come to be called, Kant claims that one does not

need any penetrating acuteness to see what I have to do in order that my volition be morally good. Inexperienced in the course of the world, incapable of being prepared for whatever might come to pass in it, I ask myself only: can you also will that your maxim become a universal law? If not, then it is to be repudiated, *and that not because of a disadvantage to you or even to others forthcoming from it* but because it cannot fit as a principle into a possible giving of universal law. (4:403, my emphasis)

As he explains in the *Groundwork* and elsewhere, Kant’s principle of moral appraisal is based not on the consequences of universalizing a maxim, but on the consistency or lack thereof in doing so. As is familiar, there are two kinds of contradiction that can arise as a result of

---

version that Kant responds to and which therefore figures most prominently in the majority of my discussion in this paper, I include a brief discussion of the egoist version not only to give thorough treatment of the allegation itself, but also in order to highlight the peculiarities of the specific version to which Kant responds.

<sup>8</sup> All references to Kant’s works cite the volume and page number of his *Gesammelte Schriften* (see Kant 1900ff.). In general, I have used the translations of Kant’s texts available in the Cambridge Edition of the Works of Immanuel Kant, and I have indicated where I have modified these translations.

universalizing a maxim and that make such a maxim immoral: a “contradiction in conception” and a “contradiction in the will.”<sup>9</sup> Thus, if Kant explicitly says that the morality of maxims is *not* decided by the consequences brought about by their universalization, one might reasonably ask what people like Mill and the others I discuss below are claiming when alleging that Kant is an implicit consequentialist.

There are two important parts to Mill’s allegation. First, Mill challenges the idea that Kant’s contradiction tests can determine the morality of maxims. As he states in the first passage quoted above, contrary to what Kant may claim, Mill believes that there is *no* logical or physical<sup>10</sup> contradiction involved in all rational beings adopting immoral rules of conduct. If, according to Mill, all rational beings are able to adopt immoral rules of conduct without any contradiction arising, then contradiction is not a plausible criterion for distinguishing moral from immoral maxims. The first part of Mill’s allegation therefore claims that Kant’s criterion of moral appraisal is not a fit method for evaluating whether candidate maxims are moral.<sup>11</sup>

---

<sup>9</sup> See 4:424 and especially O’Neill (1989, 89). It is not my purpose, nor do I have the space, to discuss here the many details of these types of contradiction and Kant’s universalization tests. For a plausible defense and a survey of the various interpretive options, see Korsgaard (1985). For a recent and very persuasive argument that there is ultimately only one kind of contradiction, see Kleingeld (2017).

<sup>10</sup> By referring to logical and physical contradiction or impossibility it is unclear what position of Kant’s Mill has in mind. The most plausible option, I take it, is that Mill is referring to Kant’s contradiction in conception test, rather than the contradiction of the will, but I have been unable to find any discussions of this in the secondary literature on Mill and *Utilitarianism*. What the secondary literature confirms, and which I discuss below, is that Mill does not interpret Kant as an egoist; as I discuss shortly Mill takes Kant to be an implicit utilitarian (see e.g. Fuchs 2006 and West 2007, pg. 103).

<sup>11</sup> In this respect the allegation of Kant’s implicit consequentialism has important similarities to the empty formalism objection. As Fabian Freyenhagen correctly points out, the empty formalism objection, at least as classically formulated by Hegel, is best captured by a cluster of related criticisms, one of which is that consistency or contradiction is not a fit test for determining whether candidate duties are genuine duties (see Freyenhagen 2011, 165).

The second part of Mill's objection claims that since Kant's conception of moral appraisal does not make sense in the way he describes, in order to make it plausible we have to interpret it along consequentialist lines. What a universalization test shows us, says Mill, is that maxims adopted on a universal scale will have good or bad consequences, and *this* feature is what would make them immoral if Kant's principle were to make sense. Mill is therefore arguing that, since contradiction is of no help in determining the morality of maxims, we must interpret Kant as an implicit consequentialist if his view is to be plausible. However, it is important to keep in mind that Mill and the others discussed below are not simply offering a misinterpretation of Kant's principle of moral appraisal and ignoring the fact that Kant explicitly denies that consequences are involved in the evaluation of maxims. The claim, rather, is that the principle of moral appraisal that Kant himself offers is implausible, and in order to make charitable sense of his position, Kant must be interpreted as an implicit consequentialist. The allegation of Kant's implicit consequentialism is therefore made in the spirit of charity, and is an attempt to make his position more plausible than he himself presents it.

What is important to note about Mill's version of the allegation is that he is not interpreting Kant egoistically. As Mill states in the quotation above, we must consider the consequences that the universal adoption of maxims would have "*with benefit to their [rational beings] collective interest*". Mill's version of the allegation is therefore, unsurprisingly, a utilitarian one. Indeed, Mill's discussion of Kant takes place in a passage where he is discussing the feeling of resentment as concerning the collective rather than individual interest. Mill believes that if our feeling of resentment is moral in character, then what we have in mind is an affront to society as a whole and the collective interest, not personal interest. Mill claims that even for "anti-utilitarian moralists", e.g. Kant, this resentment concerns the collective, rather than merely individual interest (see Mill 1969, 249).

In sum, Mill's version of the allegation states that Kant's conception of moral appraisal must be implicitly utilitarian if it is to be capable of determining the morality of maxims. Because contradiction fails to be a criterion of moral appraisal, it is the consequences that the universal adoption of a maxim would have on the collective interest that decides the morality of maxims. On Mill's view, a charitable interpretation conceives of Kant as an implicit consequentialist, and indeed an implicit utilitarian.

There is more than one version of the allegation that Kant is an implicit consequentialist, and Mill presents only the utilitarian version. Arthur Schopenhauer voiced

another version in his prize essay *Über die Grundlage der Moral* [On the Basis of Morality]. Similar in structure to Mill's version of the allegation, in this text Schopenhauer argues, first, that Kant's "*highest principle*" (2010, 155) of morals, namely the categorical imperative, is incapable of appraising the morality of maxims in the way he describes. In particular, Schopenhauer claims that if we use the categorical imperative to appraise maxims, "I can perfectly well will injustice and unkindness as a universal maxim, and regulate the world accordingly." (2010, 158) On Schopenhauer's view, then, Kant's principle of morality is not a plausible criterion of moral appraisal. Schopenhauer claims that "the fundamental law put forward by Kant is obviously not yet the moral principle itself, but at best a heuristic rule for it, i.e. an indication as to where it is to be sought." (2010, 156) According to Schopenhauer, Kant's categorical imperative actually judges the morality of maxims by means of "*egoism*":

The instruction – contained in *Kant's* highest rule – of how to find the real moral principle rests, then, on the tacit presupposition that I can will only *that* state in which I am best off. Since when I ascertain some maxim that is to be followed universally, I must necessarily consider myself not always merely as the active, but also potentially and on some occasions as the passive party; my egoism decides for justice and loving kindness from this standpoint not because it desires to *practice* them, but because it desires to *receive* them. (Schopenhauer 2010, 156)

Schopenhauer believes that although it may not be obvious on the surface, if we look "deeper in the text" (Schopenhauer 2010, 156) we find that Kant's categorical imperative is based on egoism. As his references to an agent considering themselves as the "passive party" in the above quotation indicate, Schopenhauer believes that it is primarily the consequences that the universal adoption of a particular maxim might have *for me* that decide the morality of a maxim. At the forefront of Schopenhauer's mind here is primarily Kant's contradiction of the will test, as his language in this essay makes clear.<sup>12</sup> With respect to the examples of the false promise and the maxim of beneficence, for example, Schopenhauer argues that it is "*reciprocity*" that makes such maxims morally unacceptable, i.e. that others would reciprocally make false promises to *me*, and be unwilling to assist *me*. Schopenhauer therefore believes that Kant's conception of moral appraisal is implicitly consequentialist, but in his case it is not the positive or negative consequences on the collective interest that decide the morality of maxims, rather it

---

<sup>12</sup> See e.g. "Thus only the maxim itself, of which I *can* will that all should act according to it, would be the real moral principle." (Schopenhauer 2010, 156)



is the effect that their universalization would have on individual interest.

The above are what might be the two most well-known versions of the allegation that Kant offers an implicitly consequentialist theory of moral appraisal. Numerous others have made the same allegation in the history of Kant scholarship and criticism, and it has been termed differently as well.<sup>13</sup> What has thus far been noticed very seldom in the literature, however, is the fact that this allegation was voiced in Kant's time, and that he is likely to have been aware of it. In fact, this allegation was voiced twice very soon after the publication of the *Groundwork* in 1785 and in the remainder of this section I discuss these two instances in detail.

First, Gottlob August Tittel put forward a version of the allegation in his 1786 book *Ueber Herrn Kant's Moralreform*. Like Mill, Tittel argues that the categorical imperative seems to refer to the consequences that universalizing a maxim might have for both oneself and others. After paraphrasing the categorical imperative<sup>14</sup> Tittel makes the following claim:

But does this law not immediately refer to the consequences, which such a maxim, according to its universalization, would have for myself and others? Only from the consequences can I first decide whether I can wish and will that it [my maxim] become universal or not. (Tittel 1786, 14)

The first point to mention is that Tittel's allegation is more similar in spirit to Mill's version of the allegation than to Schopenhauer's. Tittel does not take Kant to be advancing an egoist thesis but thinks it is the consequences "for myself and others" that must be considered, on Kant's view. Second, Tittel's allegation also has the same structure as the others in that he believes Kant's position is *plausible* only if it takes consequences into account. As he says in the second sentence in the passage above, it is "only" by considering the consequences that I can decide whether or not a maxim should be universalized. Indeed, Tittel stresses this point later in the text:

but this law itself, which Herr Kant believes to have discovered out of pure reason – if it is not an entirely empty and sterile expression, and should rather be capable of some application, is

---

<sup>13</sup> See e.g. Ebbinghaus (1968, 220), Simmel (1904, 97-8), Jodl (1889, 18), Broad (1965, 130), and Singer (1961, 262). The most common name for the allegation is to refer to it as the "teleological" interpretation of Kant's ethics. For a discussion, see Johnson and Cureton (2016).

<sup>14</sup> "act in such a way that you can also will that your maxim (your practical principle) can become universal" (Tittel 1786, 13). All translations of Tittel and Pistorius are my own.

entirely empirical according to its content, i.e. it refers immediately to the consequences and effects as we know them from experience. (Tittel 1786, 33)

As this passage makes clear, Tittel thinks that if consequences, and thus empirical considerations, are *not* brought into the picture, then Kant's categorical imperative has no application, i.e. cannot help decide the morality of maxims, and as such is "empty" and "sterile." Tittel therefore makes an allegation very similar to Mill, though here it is not contradiction, which Mill calls logical or physical impossibility, that Tittel identifies as problematic, but merely universalization on its own. According to Tittel, because universalizing a maxim is empty and sterile, i.e. does not help us determine the morality of our maxims, we must interpret Kant as considering the consequences that such universalization would have *for oneself and others* if his principle is to have any application.

The second instance of this allegation raised during Kant's time was advanced by Hermann Andreas Pistorius in his early review of the *Groundwork*, published in 1786 in the *Allgemeine deutsche Bibliothek*. In this review, Pistorius disagrees that a consideration of the potential consequences is not involved in Kant's principle of moral appraisal. In relation to the famous case of the maxim to make a lying promise, Pistorius claims that Kant's evaluation of this maxim "only has the intended consequences at its basis." (Pistorius 1786, 453) Pistorius challenges what principle is actually at work in our evaluation of this maxim:

is then the abolition of the reciprocal trust necessary for the sake of the lives, needs and business of human beings, the impossibility of passing things on to others via promise, not a true disadvantage contracted by me and others, which in fact must first be made aware to me via experience, and is this disadvantage not the only reason why the maxim to help oneself through lying promises does not qualify as a universal lawgiving? (Pistorius 1786, 454)

In this passage, Pistorius makes two important points. First, contrary to what Kant might say, Pistorius suggests that the maxim to make a lying promise is disqualified as a universal law because the universalization of such a maxim would have disadvantageous consequences. Second, and related to the first point, Pistorius claims that we cannot determine whether the maxim to make a lying promise is universalizable a priori and by mere thought alone. On the contrary, we need to consult experience in order to see that universalizing the maxim would be disadvantageous and therefore ought not to take place.<sup>15</sup> Pistorius's main point, however, and similar to the others discussed above, is that he does not see how the universalization of

---

<sup>15</sup> This second point will be important when we turn to Kant's response in the next section.

maxims and the contradiction (either in conception or of the will) potentially involved therein can be an indicator or test of the morality of maxims unless the consequences of universalization are taken into account. In other words, Pistorius does not understand how the contradiction in conception and contradiction in willing tests function, or at the very least does not see them as plausible ways of determining the morality of maxims. He then attributes to Kant a generic consequentialist view – that what makes a maxim morally right or wrong is the consequences that the universalization of such a maxim would have on collective interest. Indeed, Pistorius’s version of the allegation is not egoist: the disadvantage that makes a maxim wrong is “contracted by me *and others*,” and not just oneself, thus his version of the allegation is closer to Mill’s than to Schopenhauer’s. In sum, Pistorius believes that, implicit within Kant’s conception of moral appraisal is a consequentialist principle, according to which it is the consequences that universalizing a maxim would have on “me and others” that decides its morality, and that we must interpret Kant in this way in order to charitably make sense of his position.

Tittel’s and Pistorius’ versions of the allegation are important because there is reason to believe that Kant was familiar with them both, and that he read them relatively soon after their publication. With respect to Tittel, in a letter to Kant from Johann Erich Biester dated June 11, 1786 we are told that Kant wanted to write a defense “against the attacks of Herren Feder and Tittel” (10:457). Such a defense was likely planned to be included in the second *Critique*, but it has been speculated that Kant eventually abandoned this plan because of his desire to complete the third *Critique*.<sup>16</sup> Although we do not have an extended engagement with Tittel in the second *Critique*, there is at least reference to an objection both Tittel and Johann Friedrich Flatt<sup>17</sup> made to the *Groundwork*, namely that Kant does not introduce a new principle of morality, but only a new formula (see Tittel 1786, 35 and 5:8n).

The evidence that Kant read Pistorius is even more compelling. First, although Pistorius published the many reviews he wrote anonymously, his identity as the author of a review of the *Groundwork* was reported to Kant by Daniel Jenisch in a letter dated May 14, 1787, where we read: “Your reviewer in the [*Allgemeine*] *deutsche Bibliothek* apparently is Provost

---

<sup>16</sup> On this point see Beiser (1987, 184) and also Kant’s letter to Christian Gottfried Schütz, dated June 25, 1787 (10:490).

<sup>17</sup> See Flatt (1786) for his review of the *Groundwork*.

Pistorius on Femarn, the translator of Hartley: his review of your Foundations [*Grundlage*] etc., although, with all ostensible seriousness, it does not go deep enough, has found many adherents because in morality minds happen to be tuned by popularity.” (10:486-7) That Kant read and took seriously the objections Pistorius raised in his review is demonstrated by the fact that Kant explicitly responds to many of them in the second *Critique*. In the Preface to the second *Critique*, for example, Kant says the following:

In the second chapter of the Analytic I have, I hope, dealt adequately with the objection of a certain reviewer of the *Groundwork of the Metaphysics of Morals*, one who is devoted to truth and astute and therefore always worthy of respect: that *there the concept of the good was not established before the moral principle* (as, in his opinion, was necessary). (5:8-9)

The reviewer is Pistorius, and Kant is referring to what has come to be called the “priority of the good” objection.<sup>18</sup> In the final section of this paper I will return to this objection, for it is by way of responding to it, I will argue, that Kant clarifies the core difference between his own moral theory and a broadly consequentialist one. Pistorius is also one of the first to object that Kant’s moral theory is of use only in theory, and not in practice (see Pistorius 1786, pg. 463). Kant responds to this objection in chapter one of the Analytic (see 5:35). Thus it is clear that Kant was familiar with Pistorius’ review of the *Groundwork* and made a serious effort to respond to at least some of the points raised in it in the second *Critique*.

In summary, in this section I have sought to clarify the nature of the allegation of Kant’s implicit consequentialism, as well as suggest that Kant was likely aware of two early instances of the allegation, and indeed that he knew of them before publishing the second *Critique*. Since there is good reason to assume that Kant was aware of this allegation, it is plausible that Kant responded to it. In the following section I argue that Kant did indeed respond to the allegation. I reconstruct this response before drawing some conclusions about what this means for what Kant saw as the essential difference between his view and a broadly consequentialist one, a task which I accomplish in the final section of this paper.

---

<sup>18</sup> See e.g. Gesang (2007, XXVI).

## II. Kant's Response: Empiricism of Practical Reason, the Good, and Moral Appraisal

However familiar the allegations of Mill and Schopenhauer (and perhaps Tittel and Pistorius) may be to us, Kant is not thought to have offered a response to the allegation that his principle of moral appraisal is implicitly consequentialist. This is for good reason: neither in the second *Critique*, nor in any other works published after the texts by Tittel and Pistorius does Kant offer a clear and explicit response to the allegation. However, given there is good reason to believe that Kant was familiar with Tittel's and Pistorius' formulations of the allegation, which I have argued for in the previous section, it is possible and indeed likely that Kant has an indirect and/or implicit response to it. Where this response is located is a different question, but an obvious place to look is the second *Critique*, a work published relatively soon after the texts by Tittel and Pistorius appeared, and in which, as I have mentioned, it is one of Kant's goals to deal with a number of objections raised against both the first *Critique* and the *Groundwork*.<sup>19</sup> An even more specific place to look would be the section of the second *Critique* in which Kant deals with a number of the objections Pistorius raised in his review of the *Groundwork*<sup>20</sup>, namely the second chapter of the *Analytic*. I will argue that it is in this section that we find Kant's response to the allegation of his implicit consequentialism. More specifically, my claim will be that Kant responds to the allegation in two ways: first, he responds indirectly by clarifying that his own view of moral goodness and moral appraisal are distinct from a view he calls "empiricism of practical reason"; second, he responds directly by identifying the specific error a consequentialist makes, namely taking the law of nature as the actual principle of moral appraisal, rather than a mere model or 'type' of the moral law.

---

<sup>19</sup> For this point see Klemme (2010, pg. 12 – 18).

<sup>20</sup> In addition to his response to Pistorius' "priority of the good" objection (5:62-4), this chapter also contains Kant's response to Pistorius's objection to the claim (see 4:394-7) that the purpose of reason is to cultivate a good will (see Pistorius 1786, pg. 450).

## The Indirect Response

Chapter two of the *Analytic of Practical Reason* can be divided relatively neatly into three subsections.<sup>21</sup> It is only in the first subsection (5:57.17 – 65.04) that Kant deals directly with the topic described in the chapter's title, namely "the concept of an object of practical reason." The second subsection (5:65.05 – 67.23), by far the most obscure, concerns what Kant calls the "categories of freedom." The third and final subsection (6:67.24 – 71.25) has its own title and deals with the "Typic [*Typik*] of pure practical judgement". My focus in this part of the paper is what Kant discusses in the first and third sub-sections of chapter two.<sup>22</sup>

In the first subsection of chapter two Kant discusses several issues, but the focus is on the topic described in the title, namely "the concept of an object of pure practical reason" (5:57). In the first instance, objects of practical reason are the things we call good and evil: "The only objects of a practical reason are therefore those of the *good* and the *evil*." (5:58) The kinds of things that can be good and evil are, in the first instance, objects of practical reason or things that can be brought about by means of action: "To be an object [*Gegenstand*] of practical cognition so understood signifies, therefore, only the relation of the will to the action by which it [*er* – i.e. the object – M.W.] or its opposite would be made real" (5:57). However, action itself can be an object of practical reason as well, not just those things brought about by *means* of

---

<sup>21</sup> See Beck (1960, pg. 126-7) and Sala (2004, pg. 136-7).

<sup>22</sup> I focus on these two subsections not only because they are the most relevant for my purposes in this paper, but also because comparatively little literature exists on these two sections specifically, and on the second chapter of the *analytic* as a whole. For example, while most other sections of the second *Critique* receive ample discussion in the relatively recently published Cambridge Critical Guide to the text (see Reath and Timmermann 2010), there is no chapter dealing with the second chapter of the *Analytic*. The best discussions of the second chapter in general are Bacin (2001), Basaglia (2016), Beck (1960) and Sala (2004), and see also Pieper (2002). For recent work on whether Kant subscribes to the 'guise of the good' thesis, see Reath (2015) and Bacin (2019). For the second subsection on the "categories of freedom", see Bader (2009), Bobzien (1988), and Zimmermann (2011) and (2016). For the third subsection, see Westra (2016).

action.<sup>23</sup> Indeed, a few pages into the chapter Kant actually lists the things he believes can be classified as good and evil, and thus as objects of practical reason:

Thus good or evil is, strictly speaking, referred to actions, not to the person's state of feeling, and if anything is to be good or evil absolutely (and in every respect and without any further condition), or is to be held to be such, it would be only the way of acting, the maxim of the will, and consequently the acting person himself as a good or evil human being, that could be so called, but not a thing. (5:60)

An object of practical reason can therefore be actions, ways of acting, maxims, and characters, and these are the things that can be called good or evil.

This is only a very general description of the kinds of things that are good and evil, and not a criterion that helps us decide *which* actions, characters, maxims etc. are good and which are evil. When Kant turns to this question, he spends a considerable amount of time outlining and criticizing a view with which he contrasts his own. Kant dubs this view “*empiricism* of practical reason” (5:70), and he defines it as the position “which places the practical concepts of good and evil merely in experiential consequences (so-called happiness).” (ibid.) According to this alternative view, the concept of the good is equated with “the concept of something whose existence promises pleasure and thus determines the causality of the subject, that is, the faculty of desire, to produce it.” (5:58) A core element of this view, as Kant describes it, is that the concept of the good is decided by experience: “because it is impossible to see a priori which representation will be accompanied with *pleasure* and which with *displeasure*, it would be up to experience alone to make out what is immediately good or evil.” (5:58) The faculty that decides what is good and evil in this case is the feeling of pleasure and displeasure (see 5:58).

As can be seen by this definition, Kant considers it a core feature of empiricism of practical reason that good and evil are defined in terms of consequences, which he here identifies with happiness. What is important to note is that Kant does not distinguish between the consequences for individual happiness and those for collective happiness. In general, Kant understands empiricism of practical reason to be a broad category of ethical theory that encompasses a number of specific views. Among these views are at least two kinds of

---

<sup>23</sup> That Kant includes actions themselves among the things that can be classified as good and evil, and not only what can be brought about *by means of* action, fits well with his claim in the *Groundwork* that when we act morally we *take* a practical interest in the action itself (see 4:413fn).

consequentialism: both the egoist version and the utilitarian version that I discussed in section (I). When Kant criticizes empiricism of practical reason, and contrasts his own view with it, I therefore take him to be implicitly responding to both versions of the allegation by way of differentiating his own view from a variety of others, including consequentialist conceptions of ethics.<sup>24</sup>

As mentioned, Kant spends a considerable amount of time outlining and criticizing this alternative position because he wants to contrast it with his own, and therefore clarify the essential ways in which his moral theory is distinct from it. The first problem with this empiricist view, and the first way in which Kant sees it as distinct from his own, concerns its concepts of good and evil. As Kant sees it, empiricism of practical reason fails to distinguish between two closely related terms, both of which end up ambiguously falling under the category of ‘the good’: on the one hand there is what is morally good and evil (good and evil proper) and on the other there is the agreeable and the disagreeable, i.e. well-being and ill-being. The latter pair of terms signify “only a reference to our state of *agreeableness* or *disagreeableness*, of gratification or pain” (5:60), i.e. pleasure and displeasure (see also 5:58).<sup>25</sup>

---

<sup>24</sup> I want to stress here that, as a broad category of ethical theory, I think Kant intends empiricism of practical reason to be a view that encompasses more positions than those that can obviously be classified as consequentialist. One other view that might fall under the broad category of “empiricism of practical reason” is the view Kant calls “eudaimonism” (see e.g. 8:395 and 6:377 and Walschots 2017). On a basic level, a view subscribes to empiricism of practical reason if it defines good and evil in experiential terms, i.e. empirically. I think Kant would understand consequentialist theories of ethics to be a further subset of this kind of ethical theory in that they further specify that actions, rules of action, etc. are judged morally based on the consequences they would have on the things identified as good and evil, concepts which Kant thinks this view *necessarily* defines empirically. I briefly discuss why Kant thinks this in the final section of this paper. What other theories might be classified as endorsing empiricism of practical reason and where to draw the lines between them all are questions that I cannot address here.

<sup>25</sup> Kant adds here that the ambiguity of these terms is problematically preserved in Latin, but avoided in the German language: whereas Latin has only the terms *bonum* and *malum* to cover both the good and the agreeable, and the evil and the disagreeable, German is fortunate



Insofar as empiricism of practical reason defines good and evil in terms of what promises pleasure and displeasure, this view defines good and evil not only experientially (hence the qualification “empiricism” of practical reason), but it also confuses moral good with well-being, and moral evil with ill-being. For Kant, these are two distinct types of value, i.e. they are distinct in kind, they do not belong on the same continuum, and thus have no overlap or approximation.<sup>26</sup>

It is important to notice that Kant’s characterization of empiricism of practical reason never claims that it *equates* agreeableness, pleasure, or well-being (and conversely disagreeableness, displeasure, and ill-being or woe) with moral good (and evil). This is because Kant makes an important background assumption about the nature of goodness throughout the second chapter of the *Analytic* that is not immediately clear. As Kant says, “the use of language” not only requires us to distinguish between the agreeable and the good, but also “requires that good and evil always be appraised by reason and hence through concepts, which can be universally communicated, not through mere feeling, which is restricted to individual subjects and their receptivity.” (5:58) Similarly, later in the second chapter of the *Analytic* he claims that “[w]hat we are to call good must be an object of the faculty of desire in the judgement of every rational human being, and evil an object of aversion in the eyes of everyone; hence for this appraisal reason is needed.” (5:61, translation modified) The background assumption of this chapter therefore concerns the objectivity of morality: the concepts of moral good and evil are universal and necessary, i.e. they must be capable of being universally communicated, and are “necessary” (5:58) objects of the faculty of desire and aversion. For this reason, Kant believes that the only faculty capable of appraising the concepts of the morally good and evil in this sense is reason; it is only by means of reason that judgements are capable of being universally communicated among human beings, i.e. only in

---

enough to have two separate terms for these different concepts: “for *bonum* it has *das Gute* [the good] and *das Wohl* [well or well-being], for *malum* it has *das Böse* [evil] and *das Übel* (or *Weh*) [ill-being or woe].” (5:59)

<sup>26</sup> For helpful discussions of Kant’s theory of value, see Sensen (2011, esp. chapter one), and Bader (2015).

this way can such judgements be universal and necessary.<sup>27</sup>

This background assumption is important when it comes to Kant's characterization of the alternative conception of the good he discusses in the second chapter: according to empiricism of practical reason, the good is not and *cannot* be identified or equated with pleasure, because what is pleasurable and displeasurable is decided by experience and feeling. If reason is the only faculty capable of deciding what is good and evil, properly speaking, then empiricism of practical reason can only ever define good and evil, in the sense of concepts that can be universally communicated, as those things that are a *means* to the agreeable or disagreeable: "a philosopher who believed that he had to put a feeling of pleasure at the basis of his practical appraisal would have to call *good* that which is a *means* to the agreeable, and *evil* that which is a cause of disagreeableness and of pain; for, appraisal of the relation of means to ends certainly belongs to reason." (5:58) In this case, however, nothing would ever be good or evil in itself, and moral commands would only ever be hypothetical, never categorical:

although reason alone is capable of discerning the connection of means with their purposes [...] the practical maxims that would follow from the above concept of the good merely as a means would never contain as the object of the will anything good in itself, but always only good *for something*; the good would always be merely the useful, and that for which it is useful would always have to lie outside the will, in feeling. Now if the latter, as agreeable feeling, had to be distinguished from the concept of the good, then there would be nothing at all immediately good, and the good would have to be sought, instead, only in the means to something else, namely some agreeableness. (5:59)

In addition to confusing the good and the agreeable, then, another major flaw of empiricism of practical reason is that it excludes the possibility of reason judging there to be anything good (or evil) in itself. Furthermore, and as a consequence of the concepts of good and evil having this limitation, moral commands, which instruct us to seek the good and avoid the evil, would never be categorical, because categorical commands require something that is good absolutely and without exception. Empiricism of practical reason thus limits moral commands to

---

<sup>27</sup> This is of course an assumption Kant makes about the nature of morality in general and is present in all of his works on moral philosophy. Perhaps most obviously in the *Groundwork*, where Kant states that "moral laws are to hold for every rational being as such" (4:412), and indeed that the moral law must be a purely formal and rational one "if duty is not to be everywhere an empty delusion and a chimerical concept" (4:402).

hypothetical imperatives, because it would only ever command that we seek a good that is always a means to something else, namely agreeableness.

Kant's characterization of empiricism of practical reason is of course meant to highlight the ways in which his own view is distinct from it. Indeed, the features of the view that Kant criticizes can be read as features that his own view does not have. The good, for Kant, is not a mere means to what is agreeable, for this would mean there is no categorical imperative. The good is judged by reason and is universal and necessary. Most importantly, however, judgements of the good happen independently and prior to, not as a consequence of, what we find agreeable. As Kant says: "*good or evil* always signifies a reference to the *will* insofar as it is determined by the *law of reason* to make something its object; for, it is never determined directly by the object and the representation of it, but is instead a faculty of making a rule of reason the motive of an action." (5:60) The important contrast that Kant wants to make between his own position and empiricism of practical reason is that, according to his own position, it is the law of reason that first determines the object of practical reason, *via* the unconditional and categorical command it places on human beings. According to the alternative, namely empiricism of practical reason, the representation of an object first determines the will and thereby, i.e. conditionally, binds us to those actions which are a means to its realization.<sup>28</sup>

Kant not only sketches how empiricism of practical reason conceives of the good, but also offers a characterization of how this view formulates the method of moral appraisal. According to Kant, "there are two very different appraisals of an action depending upon whether we take into consideration the *good* and *evil* of it or our *well-being* and *woe* (ill-being)." (5:60) Interestingly, he does not outright reject the idea that we appraise actions (and maxims, characters, etc.) on the basis of well- and ill-being. Rather, his claim is that this is not the only, nor the primary, way in which we appraise such things:

Certainly, our well-being and woe count for a *very great deal* in the appraisal of our practical reason and, as far as our nature as sensible beings is concerned, *all* that counts is our *happiness* if this is appraised, as reason especially requires, not in terms of transitory feeling but of the influence this contingency has on our whole existence and our satisfaction with it; but happiness is not *the only thing* that counts. The human being is [...] not so completely an animal

---

<sup>28</sup> I discuss the relationship between and priority of the law and the good in more detail in the final section of this paper.

as to be indifferent to all that reason says on its own and to use reason merely as a tool for the satisfaction of his needs as a sensible being. [...] besides this he has it for a higher purpose: namely, not only to reflect upon what is good or evil in itself as well – about which only pure reason, not sensibly interested at all, can judge – but also to distinguish the latter appraisal altogether from the former and to make it the supreme condition of the former. (5:61-2)

Empiricism of practical reason, according to Kant, therefore makes the additional mistake of assuming that the only way in which we appraise actions, maxims, etc. is in terms of their relation to well-being.<sup>29</sup> Not only does this view have a distinct, and mistaken, conception of value, but its principle of moral appraisal is wrong-headed as well. Indeed, not only ought we to distinguish this conception of moral appraisal from one that appraises objects using pure reason alone, but we ought to realize that the one, i.e. the latter, pure kind of appraisal is superior to the former, and thus that nothing is empirically good that is not first determined to be morally good.

### The Direct Response

In the second chapter of the *Analytic*, Kant not only contrasts his own conception of the good with that of empiricism of practical reason but also seeks to clarify his own conception of moral appraisal in light of this alternative. He tackles this task in the third subsection of chapter two, entitled ‘Of the Type of Pure Practical Judgment.’ Practical judgement is essentially what takes place in the appraisal of actions and involves determining

---

<sup>29</sup> In the above passage Kant’s language suggests that empiricism of practical reason appraises actions in terms of their relation to our “own” happiness exclusively, but I think a more charitable reading would consider Kant to be talking about the relation that actions have to collective happiness as well. One such relation is of course the *effect* or the *consequences* that such actions etc. can have on happiness, whether individual or collective. Thus, Kant’s main issue is with neither an egoist nor a utilitarian flavour of consequentialism specifically, but rather with the *type* of moral theory they are both thought to represent, i.e. the kind that confuses goodness with well-being (whether our own or that of others) and that judges goodness in relation to this, e.g. as a means, and thus reduces moral commands to hypothetical imperatives. Both egoism and utilitarianism make these mistakes, according to Kant, and thus both are guilty of empiricism of practical reason.

whether a proposed or real action is a case of a given rule: “whether an action possible for us in sensibility is or is not a case that stands under the rule requires practical judgment, by which what is said in the rule universally (*in abstracto*) is applied to an action *in concreto*.” (5:67) The rule involved here, and thus the associated appraisal, is not necessarily moral; it is possible to appraise whether an action contributes to our overall happiness, and thus whether its maxim qualifies as a rule of skill.<sup>30</sup> In the case of moral appraisal we are concerned with *pure* practical judgment since we are appraising whether an action or maxim stands under a *pure* rule, namely the moral law.

In the case of *pure* practical judgment, however, we are faced with a problem, and one similar to the issue discussed in the schematism chapter of the first *Critique*<sup>31</sup>, namely the following:

all cases of possible actions that occur can be only empirical, that is, belong to experience and nature; hence, it seems absurd to want to find in the sensible world a case which, though as such it stands only under the law of nature, yet admits of the application to it of a law of freedom and to which there could be applied the supersensible idea of the morally good, which is to be exhibited in it *in concreto*. (5:68)

Kant’s solution to this problem is to use a “type [*Typus*]” or model<sup>32</sup> of the moral law in our pure practical judgment, one which can be applied to actions in nature, but which also contains the formal qualities of lawfulness possessed by the moral law, especially universality and necessity.<sup>33</sup> Such a law is none other than the law of nature. In the second *Critique*, Kant characterizes his “rule of judgment” (5:69) or principle of moral appraisal as follows: “ask yourself whether, if the action you propose were to take place according to a law of nature, of which you were yourself a part, you could indeed regard it as possible through your will.”

---

<sup>30</sup> Although Kant often speaks of the appraisal of “actions” in the typic section, I think Beck (see 1960, 154ff.) is correct to say that Kant’s main concern in practical judgement is with the appraisal of maxims.

<sup>31</sup> A 137/B 176ff. For a discussion of the similarities and differences between the two problems, see Beck (1960, 155f.) and Sala (2004, 153f.).

<sup>32</sup> For a discussion of the etymology and meaning of *Typik* and *Typus*, see Beck (1960, 157n) and Paton (1946, 160-161).

<sup>33</sup> Beck identifies uniformity and purposiveness as the most noteworthy qualities of the law of nature in this context (see 1960, 159 – 161).

(5:69, translation modified)

What is important for my purposes is that Kant's discussion of the *typic* contains the closest thing to a direct, though admittedly not very obvious, response to the allegation of consequentialism in his writings. The core idea of the *typic* is that the law of nature serves only as a "type" or "model" of the moral law that we are allowed to use in pure practical judgement since it shares the formal qualities of lawfulness with the moral law. It is important to always keep in mind, however, that the law of nature is not the *actual* law of pure practical judgement, but only a type or model. According to Kant, when we use the law of nature to appraise actions, we ask ourselves questions like the following:

if *everyone* permitted himself to deceive when he believed it to be to his advantage, or considered himself authorized to shorten his life as soon as he was thoroughly weary of it, or looked with complete indifference on the need of others, and if you belonged to such an order of things, would you be in it with the assent of your will? (5:69)

When we perform such thought experiments and ask ourselves such questions, we are *not* asking ourselves whether or not it would be to our advantage or whether we would desire to be part of such a world where everyone did what we propose. Kant makes this point somewhat cryptically:

Now everyone knows very well that if he permits himself to deceive secretly it does not follow that everyone else does so, or that if, unobserved, he is hard-hearted everyone would not straightaway be so toward him; accordingly, this comparison of the maxim of his actions with a universal law of nature is also not the determining ground of his will. (5:69)

Kant makes two points in this important passage. First, he stresses his main argument of the *typic*, namely that when we imagine our proposed maxim as a law of nature we should not imagine everyone immediately doing it; imagining our proposed maxim as a law of nature is rather a way for us to think of our maxim as having the formal qualities of lawfulness (at the very least universality and necessity). The law of nature is just a model we can use, our maxims do not in fact become laws of nature.

Kant's second point in the above passage follows from the first, and is that our will is not, or better *should* not, be determined by imagining that our maxim in fact becomes a law of nature. Here is where we can start to see that Kant is directly replying to the allegation of his implicit consequentialism. If we were to eventually reject a maxim like deception or hard-heartedness, our reason for doing so, and indeed our reason for not deceiving or being hard-hearted, i.e. the "determining ground of the will," does not involve considering whether a

world in which everyone did such things would be undesirable. This is the case for the reason just explained, namely that imagining such maxims as a law of nature precisely does *not* imply that everyone deceives, is hard-hearted, etc. The law of nature is merely a model we use, and we do not imagine our maxims as actual laws of nature where everyone immediately follows them. The allegation of consequentialism, however, assumes that Kant's conception of moral appraisal does precisely this, i.e. asks us to imagine a world wherein *everyone* deceives or is hard-hearted and then evaluate whether it is a world we would *desire* to live in. In essence, then, Kant believes that those who allege he is an implicit consequentialist, and indeed consequentialists themselves and perhaps anyone who subscribes to empiricism of practical reason, make a specific kind of error, namely they take what should only be a type or model of the moral (the law of nature) and make it the actual rule of practical judgement. By clarifying that his rule of practical judgement uses the law of nature only as a model, Kant thus clarifies that his principle of moral appraisal is not consequentialist.

Kant believes himself to have shown that there are various problems with empiricism of practical reason and that his alternative view, which he dubs "*rationalism of judgement*" (5:71) is superior. Kant makes this point towards the end of the typic section, where he claims the following: "This, then, as the typic of judgment, guards against [*bewahrt vor*] *empiricism* of practical reason, which places the practical concepts of good and evil merely in experiential consequences." (5:70) In the second chapter of the analytic of practical reason Kant therefore responds both indirectly and directly to the allegation that his principle of moral appraisal is implicitly consequentialist: indirectly by distinguishing his conception of moral good and moral appraisal from that held by empiricism of practical reason, and directly by arguing that the correct rule of practical judgement takes the law of nature as only the type or model of the moral law, not as the actual moral law. In these ways, I have argued, Kant considers himself to have sufficiently clarified that he is not an implicit consequentialist.

Although Kant believes his own moral theory is distinct from empiricism of practical reason in these ways, in the next section I suggest that Kant does not consider these differences to be primary or *fundamental*. On the contrary, I argue that Kant views the fundamental difference between his own view and empiricism of practical reason to be one that concerns methodology and the priority of the right over the good.

### III. Kant and Consequentialism Revisited

I began this paper with a brief summary of the literature on both sides of the debate concerning the compatibility of Kant's ethics and consequentialism. I mentioned that while a number of scholars argue that Kant's ethics is compatible with consequentialism to varying degrees, a number of others believe that the two moral theories are fundamentally distinct. In the previous sections of this paper I argued that Kant was familiar with the charge that he was offering an implicitly consequentialist conception of moral appraisal and responded to it in the second *Critique* primarily by distinguishing his own view from what he calls "*empiricism of practical reason*." There we saw that Kant considers his view to be distinct from empiricism of practical reason in at least two essential ways: (1) their respective conceptions of the good, and (2) their principles of moral appraisal. The foregoing sections of this paper therefore confirm what others have already argued<sup>34</sup>, namely that Kant's ethics and consequentialism have different conceptions of value and that they have a distinct moral principle. What I would like to argue in this final section, however, is that these features are not what Kant himself believes to be the *fundamental* difference between his own moral philosophy and a broadly consequentialist ethical theory.

Kant reveals what he believes to be the fundamental difference between his own view and not only consequentialism, but all other moral theories over the course of responding to yet another objection raised by Pistorius in his review of the *Groundwork*, one which Kant took so seriously that he explicitly and directly responded to it. As discussed briefly in section (I), Kant mentions this objection in the Preface to the second *Critique*, where he announces what he seeks to accomplish in the second chapter of the *Analytic of Practical Reason*:

In the second chapter of the *Analytic* I have, I hope, dealt adequately with the objection of a certain reviewer of the *Groundwork of the Metaphysics of Morals*, one who is devoted to truth and astute and therefore always worthy of respect: that *there the concept of the good was not established before the moral principle* (as, in his opinion, was necessary). (5:8-9)

Relatively early on in his review of the *Groundwork*, Pistorius summarizes Kant's conception of the good will as the only thing that can be good "without limitation" (4:393), and makes the following remark:

I wish the author would have found it preferable above all to discuss the general concept of

---

<sup>34</sup> See e.g. Timmermann (2005) and Rohs (1995).



that which is good, and to determine more closely what he means by this, for we clearly must come to an agreement on this before we make something of the absolute worth of a good will. I am therefore justified, first, to ask what is good in general, and what is a good will in particular? (Pistorius 1786, 449)

The main worry here is that Kant needs to define the nature of goodness before defining the good will, because otherwise the good will could be defined as one that follows any law whatsoever, rather than the moral law specifically. As Pistorius states: “is it sufficient to establish a will as the Good that it acts merely according to any kind of principle or from respect for any law, be it as it may, good or evil? – impossible, thus it must be a good principle, a good law, the following of which makes a will good.” (Pistorius 1786, 449) In order to know when a will is good, then, we need to define the nature of goodness so that we know when a will is following a good law.

Kant responds to this objection towards the end of the first subsection of chapter two of the *Analytic*. There, Kant explains what he calls “the paradox of method in a *Critique of Practical Reason*”, namely “*that the concept of good and evil must not be determined before the moral law ... but only (as was done here) after it and by means of it*” (5:63). Kant therefore clarifies that it was his intention to define the good later and after determining the nature of the moral law. This is because Kant takes defining the good first to pre-emptively decide a fundamental question that the *Critique of Practical Reason* was supposed to investigate, namely whether or not there is such a thing as pure practical reason. As he states in the first paragraph of the Preface: the *Critique of Practical Reason* “has merely to show *that there is pure practical reason*” (5:3). The second *Critique* must therefore “leave it *undecided* in the beginning whether the will has only empirical or also pure determining grounds a priori.” (5:63)

Kant explains the problem with defining the good before the law in relation to this investigation in a long passage in the second chapter of the *Analytic*:

Suppose that we wanted to begin with the concept of the good in order to derive from it laws of the will: then this concept of an object (as a good object) would at the same time supply this as the sole determining ground of the will. Now, since this concept had no practical a priori law for its standard, the criterion of good or evil could be placed in nothing other than the agreement of the object with our feeling of pleasure or displeasure... Now, since what is in keeping with the feeling of pleasure can be made out only through experience, and since the practical law is nevertheless, by hypothesis, to be based on this as its condition, the possibility of a priori practical laws would be at once excluded [...]. (5:63)

The problem with defining the good before the law is therefore that it takes a position on and decides a question at the outset, namely the precise question that the second *Critique* is meant to investigate: whether there is such a thing as *pure* practical reason, or whether there are only empirical determining grounds of the will. Furthermore, beginning with the concept of the good leaves us with only material, i.e. empirical, practical principles and leaves no room for practical laws. As Kant says earlier on in the second *Critique*, namely in §2 of chapter one, “Theorem I: “All practical principles that presuppose an *object* (matter) of the faculty of desire as the determining ground of the will are, without exception, empirical and can furnish no practical laws.” (5:21)<sup>35</sup> If we are only left with material practical principles, as would be the case if we defined the good first, the possibility of a priori practical laws would be excluded at the outset. We need to leave it an open question whether pure reason is practical, and whether there are a priori practical laws, and this is what Kant seeks to accomplish by *not* defining the good first.

What is important for my purposes about Kant’s paradox of method and his answer to the priority of the good objection is a remark that he adds towards the end of the section, which he says “concerns only the method of ultimate moral investigation.” (5:64) Kant claims that this remark “is important” because “[i]t explains at once the occasioning ground of *all* the errors of philosophers with respect to the supreme principle of morals.” (5:64, my emphasis) Note here that Kant takes this methodological difference to be the distinguishing feature between not only his own moral theory and empiricism of practical reason (and therefore consequentialist theories of morality), but *all* other supreme principles of morality. The mistake that Kant thinks all other moralists make is the following: “they sought an object of the will in order to make it into the matter and the ground of a law [...] whereas they should first have searched for a law that determined the will a priori and immediately, and only then determined the object conformable to the will.” (5:64) To emphasize, as a mistake that Kant believes all other moralists to have made, this is an error committed by empiricism of practical reason as

---

<sup>35</sup> It should be emphasized here that Kant distinguishes his view from empiricism of practical reason *implicitly* at various points throughout the second *Critique*, both before and after the second chapter of the *Analytic*. He only *explicitly* highlights the differences between empiricism and his own view in the second chapter, however, which is why this section has been my main focus in this paper. I thank an anonymous reviewer for encouraging me to clarify this point.

well.<sup>36</sup> Indeed, the key point in relation to my argument in this paper is that Kant takes empiricism of practical reason to have the conception of goodness that it does, as well as its corresponding conception of moral appraisal, *because* it makes the methodological error of defining the good first. Kant makes this point early on in chapter two of the analytic: “If the concept of the good is not to be derived from an antecedent practical law but, instead, is to serve as its basis, it can be *only* the concept of something whose existence promises pleasure and thus determines the causality of the subject, that is, the faculty of desire, to produce it.” (5:58, my emphasis) This reveals what, in Kant’s view, is the *fundamental* difference between his own moral philosophy and a view according to which actions, maxims, etc. are judged on the basis of the consequences they have on happiness. While Kant’s own view and that of empiricism of practical reason certainly differ with respect to their conceptions of the good and moral appraisal, the above remark of Kant’s reveals that the empiricist of practical reason holds these views because of a more fundamental, methodological commitment, and which Kant regards as an error, namely the belief that the good needs to be defined before the law. As Kant says, if we define the good before the law, the good can *only* be an object that promises pleasure, and moral appraisal would then utilize this conception of goodness in turn. Kant takes his own view to be distinct in that he defines the moral law *before* determining the nature of moral goodness. It is by this means that Kant leaves it open whether or not empiricism of practical reason is the only option, and whether there might also be such a thing as pure practical reason. Thus Kant takes the fundamental difference between his own view and the empiricist of practical reason, and therefore the consequentialist as well, to be one that concerns methodology: the consequentialist makes the error of defining the good before the law, whereas Kant correctly defines the law first. This is the most fundamental difference between the two views, according to Kant, and this results in them possessing different conceptions of value and moral appraisal.

---

<sup>36</sup> Not all moral theories other than Kant’s own can be classified as subscribing to empiricism of practical reason. In addition to empiricism of practical reason and “rationalism of judgment,” there is also “mysticism of practical reason” which Kant claims “puts under the application of moral concepts real but not sensible intuitions.” (5:70-1)

## Conclusion

There are further questions to explore here. For example, one might ask whether it is true that Kant's moral theory is distinct from *all* other approaches, and in the way he suggests.<sup>37</sup> One might also wonder whether defining the good before the right does in fact commit one to defining goodness in terms of pleasure, or, similarly, one might ask whether it is possible to define the law before the good, as Kant does, yet still be a consequentialist.<sup>38</sup> Even more importantly, it needs clarifying why Kant thinks that defining the good before the right necessarily commits one to an experiential conception of the good.<sup>39</sup> I cannot answer these questions here, and merely submit them as the next steps to be taken in the discussion.

It was my aim in this paper to outline Kant's response to what I have called the allegation of consequentialism, and, in light of this, to give an account of what I consider Kant took to be the primary difference between his own view and a broadly consequentialist one. In doing so I hope to have brought to light Kant's take on a matter, which, in the past, scholars have thought him *not* to possess, namely the question of the fundamental difference between his own view and a broadly consequentialist one. On an exegetical level, I hope to have clarified some core passages in the second chapter of the *Analytic of the Critique of Practical Reason*. I also hope, however, to have contributed to a debate of substantial philosophical interest, namely the question of what, on a systematic level, separates Kantian ethics from

---

<sup>37</sup> For a recent defence of this, see Sensen (2015).

<sup>38</sup> For such an argument see Forschler (2013).

<sup>39</sup> Such a clarification would require an extended discussion that goes beyond the scope of this paper. One possibility, however, is that Kant holds such a view due to the nature of the faculty of desire and the laws associated with the higher and lower instances of it. The lower faculty of desire is governed by the law of nature, i.e. the law of what *does* happen. The higher faculty of desire is governed by the moral law, i.e. the law of what *ought* to happen. If our first methodological step is to define goodness, it is possible that we would be forced to define goodness in terms of what we already in fact *do* find good, namely pleasure, as opposed to what we *ought* to find (supremely) good, namely acting morally.

consequentialism. I have argued that what distinguishes them most fundamentally is, at least on Kant's view, a methodological priority given to either the right or the good.<sup>40</sup>

### Works Cited

- Bacin, Stefano. 2001. "Die Lehre vom Begriff des Guten in der "Kritik der praktischen Vernunft"" in *Kant und die Berliner Aufklärung. Akten des 9. Internationalen Kant-Kongresses* edited by Volker Gerhardt, Rolf-Peter Horstmann, Ralph Schumacher. Berlin–New York, Vol. III, pg. 131-140.
- 2019. "'Under the Guise of the Good": Kant and a Tenet of Moral Rationalism" in Violetta Waibel et al. (eds.), *Natur und Freiheit. Akten des 12. Internationalen Kant-Kongresses*, Berlin–Boston, Vol. III, pg. 1705-1714.
- Bader, Ralf M. 2015. "Kantian Axiology and the Dualism of Practical Reason" in *The Oxford Handbook of Value Theory* edited by Iwao Hirose and Jonas Olson. Oxford, pg. 175 – 204.
- 2009. "Kant and the Categories of Freedom." *British Journal for the History of Philosophy*. Vol. 17, No. 4, pg. 799 – 820.
- Basaglia, Federica. 2016. "The Highest Good and the Notion of the Good as Object of Pure Practical Reason" in *The Highest Good in Kant's Philosophy* edited by Thomas Höwing. Berlin, pg. 17 – 32.
- Beck, Lewis White. 1960 *A Commentary on Kant's Critique of Practical Reason*. Chicago.

---

<sup>40</sup> I would like to thank audiences at the University of Western Ontario, the University of St. Andrews, and the University of Southampton for their comments on earlier versions of this paper, as well as Corey Dyck, Jens Timmermann and two anonymous referees for their helpful feedback. I also wish to thank Trevor Bieber, Anthony Machum, Florian Marwede, and my wife Kacy Walschots for carefully proofreading the text. The research for this paper was generously supported by the Social Sciences and Humanities Research Council of Canada.

- Beiser, Frederick C. 1987. *The Fate of Reason*. Cambridge, Mass.
- Bobzien, Susanne. 1988. "Die Kategorien der Freiheit bei Kant" in *Kant: Analysen – Probleme – Kritik*. Edited by H. Oberer and G. Seel, Königshausen and Neumann, pg. 193 – 220.
- Broad, C.D. 1965. *Five Types of Ethical Theory*. Totowa, NJ.
- Cummiskey, David. 1996. *Kantian Consequentialism*. Oxford.
- Ebbinghaus, Julius. 1968. "Interpretation and Misinterpretation of the Categorical Imperative." in *Kant: A Collection of Essays* Edited by Robert Paul Wolff. London.
- Flatt, Johann Friedrich. 1786. Grundlegung der Metaphysik der Sitten. *Tübingsche gelehrte Anzeigen*, 14th piece, 16 Feb., pg. 105-112.
- Forschler, Scott. 2013. "Kantian and Consequentialist Ethics: The Gap Can Be Bridged." *Metaphilosophy*. Vol. 44, No. 1-2, pg. 88-104.
- Freyenhagen, Fabian. 2011. "Empty, Useless, and Dangerous? Recent Kantian Replies to the Empty Formalism Objection." *Bulletin of the Hegel Society of Great Britain*. Vol. 63, pg. 163 – 186.
- Fuchs, Alan E. 2006. "Mill's Theory of Morally Correct Action" in *The Blackwell Guide to Mill's Utilitarianism*. Edited by Henry R West. Oxford. pg. 139-159.
- Gesang, Bernward. Ed. 2007. *Kants vergessener Rezensent: Die Kritik der theoretischen und praktischen Philosophie Kants in fünf frühen Rezensionen von Herman Andreas Pistorius*. Hamburg.
- Hare, R. M. 1997. *Sorting Out Ethics*. Oxford.
- Heydt, Colin. 2014. "Utilitarianism before Bentham" in *The Cambridge Companion to Utilitarianism*. Edited by Ben Eggleston and Dale Miller. Cambridge, pg. 16 – 37.

- Hruschka, Joachim. 1991. "The Greatest Happiness Principle and Other Early German Anticipations of Utilitarian Theory." *Utilitas*. Vol. 3, Issue 2, pg. 165-177.
- Johnson, Robert and Cureton, Adam. 2016. "Kant's Moral Philosophy" in *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/kant-moral/> Accessed Nov. 13<sup>th</sup>, 2017.
- Jodl, Friedrich. 1889. *Die Geschichte der Ethik*. Vol II. Stuttgart.
- Kant, Immanuel. 1900ff. *Gesammelte Schriften* (Akademie Ausgabe). Vol. 1 – 22 Ed. Preussische Akademie der Wissenschaften, Vol. 23 Ed. Deutsche Akademie der Wissenschaften zu Berlin, Vol. 24ff. Ed. Akademie der Wissenschaften zu Göttingen. Berlin.
- Kleingeld, Pauline. 2017. "Contradiction and Kant's Formula of Universal Law." *Kant-Studien*. Vol. 108, No. 1, pg. 89 – 115.
- Klemme, Heiner F. 2010. "The Origin and Aim of Kant's Critique of Practical Reason." In *Kant's Critique of Practical Reason: A Critical Guide*. Ed. Reath and Timmerman. Cambridge, pg. 11 – 30.
- Korsgaard, Christine M. 1985. "Kant's Formula of Universal Law." *Pacific Philosophical Quarterly*. Vol. 66 No. 1-2, pg. 24-47.
- Mill, John Stuart. 1969. *Collected Works of John Stuart Mill Volume X: Essays on Ethics, Religion and Society*. Edited by J.M. Robson. Toronto.
- O'Neill, Onora. 1989. *Constructions of Reason*. Cambridge.
- Parfit, Derek. 2011. *On What Matters: Volume III*. Oxford.
- Paton, H.J. 1946. *The Categorical Imperative*. Essex.

- Pieper, Annemarie. 2002. "Zweites Hauptstück (57-71)" in *Immanuel Kant: Kritik der praktischen Vernunft*. Ed. Otfried Höffe. Berlin, pg. 115-134.
- Pistorius, Herman Andreas. 1786. "Grundlegung zur Metaphysik der Sitten von Immanuel Kant. Riga, bey Hartknoch. 1785. 8. 128 S." *Allgemeine deutsche Bibliothek* Bd. 22, St. 2, pg. 447 – 463.
- Reath, Andrews. 2015. "Did Kant Hold that Rational Volition is *Sub Ratione Boni*?" in Mark Timmons and Robert M. Johnson Eds. *Reason, Value, and Respect: Kantian Themes from the Philosophy of Thomas E. Hill, Jr.* Oxford, pg. 232 – 255.
- Reath, Andrews and Timmermann, Jens Ed. 2010. *Kant's Critique of Practical Reason: A Critical Guide*. Cambridge.
- Rohs, Peter. 1995. "Warum Kant kein Utilitarist war" in *Zum moralischen Denken*. Vol. 2. Edited by Fehige, Christoph and Meggle, Georg. Frankfurt, pg. 35 – 41.
- Sala, Giovanni B. 2004. *Kants "Kritik der praktischen Vernunft": Ein Kommentar*. Darmstadt.
- Scheffler, Samuel Ed. 1988. *Consequentialism and its Critics*. Oxford, 1988.
- Schopenhauer, Arthur. 2010. *The Two Fundamental Problems of Ethics*. Cambridge, 2010.
- Sensen, Oliver. "2015. Moral Obligation and Free Will" in *Kant's Lectures on Ethics: A Critical Guide*. Edited by Lara Denis and Oliver Sensen. Cambridge, pg. 138 – 155.  
- 2011. *Kant on Human Dignity*. Berlin.
- Simmel, Georg. 1904. *Kant*. Leipzig.
- Singer, Marcus. 1961. *Generalization in Ethics*. New York.



- Timmermann, Jens. 2015. "What's Wrong with Deontology?" in *Proceedings of the Aristotelian Society*. Vol. 115, No. 1 pt. 1, pg. 75-92.
- 2014. "Kantian Ethics and Utilitarianism" in *The Cambridge Companion to Utilitarianism*. Edited by Ben Eggleston and Dale Miller. Cambridge, pg. 239 – 257.
- "2005. Why Kant Could not Have Been a Utilitarian." *Utilitas*. Vol. 17, No. 3., pg. 243-264.
- Tittel, Gottlob August. 1786. *Ueber Herrn Kant's Moralreform*. Frankfurt and Leipzig.
- Walschots, Michael. 2017. 'Kant on Moral Satisfaction.' *Kantian Review*. Vol. 22, No. 2, pg. 281-303.
- West, Henry R. 2007. *Mill's Utilitarianism: A Reader's Guide*. London.
- Westra, Adam. 2016. *The Typic in Kant's Critique of Practical Reason*. Berlin.
- Zimmermann, Stephan. 2011. *Kants 'Kategorien der Freiheit'*. Berlin.
- Ed. 2016. *Die "Kategorien der Freiheit" in Kants praktischer Philosophie: Historisch-systematische Beiträge*. Berlin.