

Jens Timmermann

The Law and the Good. Kant's Paradox of Method

*There is nothing
Either good or bad but thinking makes it so.*
William Shakespeare, *Hamlet*, II.ii

1

In Chapter II of the *Analytic of the Critique of Practical Reason*, Kant sets out to defend the priority of the moral law over the good. It is a thesis that a casual reader of the *Groundwork of the Metaphysics of Morals* might just overlook. The *Groundwork* is a book about the good will, and goodness plays a central part from page one. In that sense, goodness takes precedence. In Section I, however, we learn what makes the good will good: its being determined by a formal law, encountered by human beings in the guise of an unconditional command. It is a law that we are bound to respect, a law that restricts what we can rationally want. It defines a kind of goodness that is radically different from the value of the objects we like. Section I culminates in the question

what kind of law can that possibly be, the representation of which – even without regard for the effect expected from it – must determine the will for it to be called good absolutely and without limitation?¹

The answer given a few lines further down is: the principle of universal willing, dubbed the ‘categorical imperative’ a little later, in Section II. Kant returns to the topic at the end of that section. “We can now end”, he writes, “where we set out from at the beginning, namely with the concept of an unconditionally good will.”² In what follows, he reviews all versions of the categorical imperative as formulations of the principle of good willing. On closer

¹ Kant: GMS, AA 04: 402.01–04. All translations broadly follow the standards set by the Cambridge Edition but are essentially my own.

² Kant: GMS, AA 04: 437.06–07.

Jens Timmermann, University of St Andrews, jt28@st-andrews.ac.uk

<https://doi.org/10.1515/9783110467888-047>

inspection, then, the moral law takes precedence in the *Groundwork* because it defines what is morally good. For Kant, form determines matter.³

2

This point was not lost on Hermann Andreas Pistorius, the anonymous reviewer of the *Groundwork* for the *Allgemeine deutsche Bibliothek*.⁴ Pistorius passionately disagrees with Kant by arguing for the priority of the good. He should have wished, he says, that Kant “had deigned to discuss before all other things the general concept of what is *good*”. Pistorius is not content with the idea that the good is “what is universally endorsed and esteemed”.⁵ Appeal to general agreement is unsatisfactory because it just leads to the question “why it is endorsed and esteemed, and whether justly and reasonably or not”.⁶ Like many readers after him, Pistorius is struggling to make sense of a purely formal foundation of value. He fails to see, he says,

how one can assume that anything is entirely and absolutely good, or how one can call anything good, that is in actual fact good for nothing, and no more how there could be an assumption of a will that is good absolutely, considered merely in itself.⁷

3 This claim is not often heard in contemporary Kantian ethics, which tends to treat a value – such as rational nature, freedom or dignity – as fundamental. Unequivocal endorsements of Kant’s formalism (like Johnson, 2007 and Sensen, 2011) are the exception. Cf. Johnson, Robert: *Value and Autonomy in Kantian Ethics*. In: *Oxford Studies in Metaethics* 2 (2007), 133–148; Sensen, Oliver: *Kant on Human Dignity*. Berlin/Boston 2011.

4 Pistorius’ review was published in the 1786 volume of *Allgemeine deutsche Bibliothek*, 447–463 and is here quoted, in my own translation, as ‘AdB’. The text can also be found in Landau, 1991 and in Gesang, 2007. Cf. Gesang, Bernward: *Kants vergessener Rezensent*. Hamburg 2007; and: *Rezensionen zur Kantischen Philosophie, 1781–87*. Ed. A Landau. Bebra 1991. For a summary of Pistorius’ criticisms of the *Groundwork* see Beiser, Frederick C.: *The Fate of Reason: German Philosophy from Kant to Fichte*. Cambridge, MA, and London 1987, 190–192. However, Beiser curiously misjudges the force of Kant’s reply to Pistorius regarding the priority of the law or the good, which goes far beyond re-iterating his formalist project.

5 Pistorius: Adb 449 (See footnote 4.)

6 Pistorius: Adb 449.

7 Pistorius: Adb 449. – Note that the sentence that follows is corrupt. Summarising Kant’s argument, Pistorius needs to say that “the will is meant to be absolutely good *just* with reference to its principle or a law for the sake of which it acts, *not* with reference to any of its objects” (my emphasis), not vice versa, as the printed text and all editions have it.

If the will is to be good with reference to its principle, “does it suffice to make a good will good”, Pistorius asks, “that it act just according to some principle or from respect for some law, be it what it may, good or evil? – Impossible!” The principle of the good will cannot be just any old principle. It must be a *good* principle to command authority, “and thus the question of what is good returns once again” (AdB 449).⁸ What makes it the case that we should obey a principle? In other words, what makes a good principle good? So, even if we had pushed the question “back from the will to the law”, we would

now have to give a more satisfactory answer; i.e. now we must finally resort to some object or the final end of the law, and must turn to what is material for help, because we cannot make do with what is formal about either the will or the law.⁹

In short, we have no reason to follow a law unless it is good. But what makes a good law good is that it aims at a good end or object. For Pistorius, form follows matter.¹⁰

8 Gesang underestimates the force of this argument when he says that Pistorius “made the mistake” to think it was Kant’s intention to distinguish several laws, good and bad, and to raise an additional question as to which of these laws the human will should enact. See: Gesang, Bernhard: *Kants vergessener Rezensent*. Hamburg 2007, XXVIII. – Pistorius does not attribute a multiplicity of possible practical laws to Kant. This is part of his challenge. Pistorius raises the fundamental question why we should act on a law if, on the face of it, there is nothing to recommend it. This is the question of normativity, popularised in more recent times by philosophers who call for a ‘reason’ to act on the moral law even though it can be shown that it applies to the human will directly. See: Herman, Barbara: *Leaving Deontology Behind*. In: *The Practice of Moral Judgment*. Cambridge, MA, and London 1993, 208–240. Pistorius’ point is thus a version of the *Why be moral?* question, which ‘constructivist’ or ‘constitutivist’ reconstructions of Kant’s ethics seek to escape by arguing that the moral law is the law that naturally and inescapably governs rational willing.

9 Pistorius: AdB 449.

10 Note that Section I of the *Groundwork* is not a straightforward analysis of good willing. Kant explicitly dismisses that strategy at AA 04: 397.01–10, where he shifts focus to acting from duty (i.e. to good willing under the restrictions and limitations faced by human beings). See Timmermann, Jens: *Kant’s “Groundwork of the Metaphysics of Morals”: A Commentary*. Cambridge 2007, 25f.

3

This objection would, if sound, undermine the foundations of Kantian ethics. That is why Kant decided to provide a detailed reply in the second *Critique*. The topic is so close to his heart that he trails his response in the Preface:

In the second chapter of the *Analytic I* have, I hope, dealt adequately with the objection of a certain truth-loving and astute, yet at the same time always respect-worthy reviewer of the *Groundwork of the Metaphysics of Morals*: that *there the concept of the good was not established before the moral principle* (as, in his opinion, was necessary).¹¹

It is clear from this and from the actual argument we are about to discuss that Kant accepts Pistorius' account of his formalist ethics as essentially correct. A formal law, independent of content, must determine the will for it to be morally good. The *Critique of Practical Reason* departs from the *Groundwork* in that it no longer starts with the notion of the good will but with a discussion of practical principles. Goodness – the goodness of objects of practical reason – is mentioned only some sixty Academy pages into the book.¹² But this is a change in presentation, not in substance.¹³ The good will is reintroduced in the Chapter II of the *Analytic* and characterised in much the same terms as in the famous opening paragraph of the *Groundwork*.¹⁴

So, the *Critique of Practical Reason* starts with principles and moves on to the concepts of good and evil – and then the feeling of respect for this law – only once the principle of pure practical reason has been established. This is not just a departure from the architectonic of the *Critique of Pure Reason*¹⁵ and the *Groundwork*. Kant actually concedes that the overall structure of the *Analytic* must seem an unusual way to approach the subject matter of ethics. It is the 'paradox of method'

¹¹ Kant: KpV, AA 05: 08.25–09.03.

¹² The structure of the second *Critique*, which openly puts form before matter, can be seen as reacting to Pistorius' intervention.

¹³ Guyer's suspicion that the *Critique of Practical Reason* represents a fundamental change of strategy is unfounded (Guyer, Paul (2000). *Kant on Freedom, Law, and Happiness*. Cambridge, 139; cf. Reath, Andrews (2003). *Value and Law in Kant's Moral Theory*. In: *Ethics* 114, 127–155, here: 143).

¹⁴ See Kant: KpV, AA 05: 62.09–35.

¹⁵ Cf. Kant: KpV, AA 05: 16.22–33 and AA 05: 89.33–90.09.

that the concept of good and evil must not be determined before the moral law (for which, as it would seem, this concept would have to be made the foundation) but only (as was done here) after it and by means of it.¹⁶

This paradox is not a paradox in our modern, logical sense, but in the older sense of an assertion that is unexpected or contrary to received opinion. Goodness is the basic concept of practical philosophy.¹⁷ For Kant, it is a tautology to say that one has to do what is good. Frederick Beiser is therefore right to note that Kant's thesis of the priority of the law "must appear paradoxical in light of conventional moral theory".¹⁸ Moreover, goodness is fundamental in that we intuitively or naively judge actions or objects to be good before we start asking questions about what makes them good (see Brandt, 1995: 207). If so, arguing for the primacy of the law over the good is nothing less than a second Copernican revolution in philosophy (Silber, 1959).

4

But what does Kant mean when he uses the words 'good' (*gut*) and 'evil' (*böse*)? The first thing to note is that both notions¹⁹ are practical as well as objective:

What we are to call good must be an object of the desiderative faculty in the judgement of every rational human being, and what is evil an object of loathing in the eyes of everyone; and hence judging like that requires reason, in addition to sense.²⁰

Good and evil are values that can spark rational debate. When I judge something to be good – rather than, say, delicious – I expect you to agree. If you do not, I can try to persuade you by reasoned argument. The two concepts denote

16 Kant: KpV, AA 05: 62.37–63.04, emphasis removed.

17 That is why Kant uses (an interpretation of the notion of) the 'highest good' as his starting point in the *Groundwork*, see Timmermann, Jens: *Kant's "Groundwork of the Metaphysics of Morals": A Commentary*. Cambridge 2007, 15.

18 Beiser, Frederick C.: *The Fate of Reason. German Philosophy from Kant to Fichte*. Cambridge, MA, and London 1987, 190.

19 At the very least, this is true of the two terms in their proper terminological use. Kant is happy to admit that, in everyday language, we sometimes call 'good' what is agreeable or pleasant, i.e. something that is practically relevant but *not* objective. Restricting goodness to what is objectively, rationally desirable is thus an attempt to tidy up ordinary language for philosophical reasons.

20 Kant: KpV, AA 05: 60.37–61.03.

values that are essentially shareable ('impartible' or 'communicable', *mitteilbar*) because they are grounded in a faculty we all have in common. Moreover, like the agreeable but unlike beauty, good and evil are practical predicates in that they are bound up with an interest, i. e. with the possibility of volition and action. Judgement as such is not practical. We need a motive to be able to act.²¹

What is equally important: both 'good' and 'evil' must be construed in a broad sense not restricted to the moral realm. This is far from obvious. In modern German, the word for 'evil' (*böse*) is most commonly understood as referring to wickedness, malice or vice.²² For Kant, however, it is just a catch-all word for whatever has objective practical disvalue, whether moral, non-moral or instrumental. In English too, 'evil' loses its moral overtones when, in discussions of the 'problem of evil', we talk about 'natural' evil as opposed to 'moral' evil.²³

21 Good and evil are therefore much narrower than our contemporary concept of a 'value'. Having said that, it is difficult to see how, in Kant's axiology, any (positive) value that is both objective and practical would not be called 'good'.

22 According to Adelung's *Grammatisch-kritischem Wörterbuch der hochdeutschen Mundart* (1811), *böse* is "that which in all its meanings expresses the opposite of *gut*, and generally means everything that is contrary to the purposes of a rational being". The dictionary explicitly notes the distinction between the moral and the non-moral sense of *böse*; and for the latter he lists the following examples: *ein böses Auge* is a bad eye; *eine böse Nuß* is a deaf – i. e. empty – nut; and *böses Geld* is debased or counterfeit coinage. Kant himself talks about *böses Wetter* (bad weather) where he mentions the human tendency to talk about the weather (Anth, AA 07: 176 fn.); and in his drafts for the *Prolegomena* he says: "So far I have said only good things about my work. But now I must also say something bad [*etwas böses*] about it" (VAProl, AA 23: 59.22f). This is his way of apologising for the stylistic and presentational shortcomings of the book. Finally, the non-moral sense of *böse* is not altogether extinct. We still call an accident, a wound or an experience *böse* if we wish to convey that they are particularly dangerous or serious – but not necessarily morally wicked!

23 That Kant is using *böse* in its broad, non-moral meaning makes all the more sense once we realise that the contrast between *gut* and *schlecht*, so familiar to German speakers today, had not yet been firmly established in the late eighteenth century: *schlecht* was not Kant's standard word for 'bad'. In the *Critique of Practical Reason* it occurs only once, and it is clearly used in an earlier sense: someone thinks "that in case of future need he can make do with little [*sich ... schlecht behelfen zu können*]" (KpV, AA 05: 20.36). The word originally meant 'straight', as opposed to 'crooked', and then came to mean 'smooth', 'plain' (also in its pejorative sense), 'light' and, finally, 'bad'.

5

Lewis White Beck raises the question of what exactly qualifies as an 'object', i.e. what it is that Pistorius and Kant alike wish to call 'good' or 'evil'. The word is naturally taken to refer to some state of affairs an action aims to bring about. Beck correctly notes that Chapter II of the *Analytic* calls for a broader conception of 'object'. It needs to cover both "states of affairs produced by action"²⁴ and "action itself".²⁵ After all, this stage of the second *Critique* concerns practical value terms generally, which rules out a notion of 'object' that is either narrowly moral or restricted to prudential outcomes. This is in line with the typology of action familiar from the *Groundwork*.²⁶ The object differs, depending on whether an action is determined by a moral or non-moral laws. Moral action contains its primary end in itself. That is why moral action is unconditionally good, regardless of its intended or actual effects. By contrast, non-moral – i.e., in the final consequence, prudential – action is done for the sake of the intended effect. Non-moral action is good, in a limited way, if and only if it serves its purpose: to make the agent happy. If an action that aims at happiness does not succeed in making the agent happy he has done the wrong thing.²⁷

However, the 'object' in the argument against Pistorius²⁸ is, inevitably, an object distinct from the action, namely its external purpose. The idea that an action could be its own object is off the table as soon as Kant starts entertaining Pistorius' thesis that the object is primary. Pistorius' object must be an object in the narrow sense: the end chosen for its intended effect, as it affects the agent. What is more, in reply Kant raises the question how the will should connect with its – external – object. If the action were itself good, it would have to be directed by the law itself. Action and law would be intimately linked; no further connection would be necessary.

24 Beck should probably have said 'intentionally' or 'voluntarily' produced, but that is a minor quibble.

25 Beck, Lewis White (1960): *A Commentary on Kant's Critique of Practical Reason*. Chicago, 129.

26 For a comprehensive discussion of the two modes of willing see Timmermann, Jens: *Acting from Duty: Inclination, Reason and Moral Worth*. In: *Kant's "Groundwork of the Metaphysics of Morals": A Critical Guide*. Ed. Jens Timmermann. Cambridge 2009, 45–62.

27 Accordingly, Kant distinguishes two kinds of interest: interest in action, and interest in the object of an action. The former is possible only if there is a law of pure reason that determines the will directly (not a law of nature that instrumental reason uses to determine the will, to bring about an object). See Kant: GMS, AA 04: 413–414 fn. and AA 04: 459–460 fn.

28 Kant: KpV, AA 05: 58–59.

6

Let us finally turn to Kant's actual response in Chapter II of the *Analytic*. His strategy turns on a *reductio* of Pistorius' position that throughout what is good determines the law on which we ought to act. If successful, Kant can claim victory and reassert that in moral matters the law determines the good.²⁹

Let us assume, then, that the good underlies the practical law as its foundation. If so, the concept of the good can, according to Kant, "only be the concept of something whose existence promises pleasure and thus determines the causality of the subject to produce it", i.e. it thus determines the human will.³⁰ This claim is surprising because it limits our philosophical options and puts us on the path to settling Pistorius with hedonism, an unlikely foundation of objective, impartial moral norms by anyone's standards. So, why should our notion of goodness depend on the prospect of pleasure if we assume that goodness is prior to and independent of a practical law?

Kant can point to two relevant considerations. The first draws on the ordinary use of language. The agreeable, Kant says in his drafts for *Theory and Practice*, "is commonly also (relatively) called good (for the inclination of this or that human being) but is essentially different from that which is by itself and unconditionally good and is not an element of happiness at all".³¹ There appear to be two distinct divisions of goodness, "and no balance can be struck between the two".³² We have already seen that Kant's stance in the *Critique of Practical Reason* (and elsewhere) is even more radical. The 'relative' sense of good does not deserve the name of goodness at all. It is a misuse of language to call the agreeable 'good'. Still, the ambiguity might lie be-

29 Kant knows that indirect arguments have their drawbacks. According to the Doctrine of Method of the first *Critique*, it is a flaw of apagogic proofs that they do not reveal the reason *why* a position is correct – only *that* it is, on the assumption that the opponent is wrong and there are only two options. They are "more of an emergency aid than a procedure that satisfies all the aims of reason" (KrV, A 790/B 818). Here in the second *Critique*, Kant probably assumes that the constructive argument for his own position is contained in earlier parts of the *Analytic* and in the *Groundwork*. Moreover, the indirect argument against Pistorius rests on assumptions he takes both philosophers to share, e.g. that there are valid moral imperatives and that ethics must have something to say about goodness as well as laws, that there are only two ways of relating them to each other etc. Philosophically, none of this can be taken for granted.

30 Kant: KpV, AA 05: 58.12–14.

31 Kant: AA 23: 131.17–20.

32 Kant: AA 23: 131.27–28.

hind the move of settling Pistorius, who ruled out one kind of goodness as unintelligible, with the other. Unfortunately, this argument from linguistic usage is not very strong. An opponent might still reply that he does not intend to defend a conception of goodness that comes down to the agent's happiness, pleasure or well-being.

Secondly, the priority of goodness might lead to hedonism not on linguistic but on philosophical grounds. It is significant that Kant follows Pistorius in speaking of that which is good in terms of an 'object', to be realised by the human will. Now, I can picture all manner of things. But that does not, as such, make me want to obtain, create, promote, destroy, shed or shun them. They are not, as such, objects of my will. The mere representation of something is practically irrelevant unless there is a suitable connection that bridges the 'gap' between will and object.³³ If there is no law that determines the will directly, independently of a conception of the good, how does the will acquire a practical attitude, positive or negative, towards its object? Kant's answer is that this can only be done if the agent relates the thing in question to his capacity to feel pleasure and pain. We think of one object as pleasing and seek to attain it as a result; we think of another object as disagreeable, and as a consequence we seek to avoid it. Besides the law, only inclination can be practical, and inclination aims at objects that promise pleasure.

7

In what follows, Kant reiterates his well-known view that it is impossible to make out a priori "which representation will be accompanied with *pleasure* and which with *displeasure*". So, on the view attributed to Pistorius, the agent would have to rely on experience to determine which objects are "immediately good or evil".³⁴ We do not know whether we will like something

³³ Note that it is taken for granted that the concept of goodness is not contemplative but practical. A good object cannot be an object of quiet admiration; in that the good differs from beauty. It's impossible to say: yes, it's good (in its own right, not just instrumentally) but I have no interest in pursuing it. Put slightly differently, 'good' and 'evil' are the – only! – concepts that concern objects of practical reason, cf. KpV, AA 05: 58.06–07.

³⁴ Kant: KpV, AA 05: 58.15–17. – In the *Groundwork*, Kant similarly argues that "the elements that belong to the concept of happiness are one and all empirical, i.e. must be borrowed from experience" (GMS, AA 04: 418.05–07). Of course, 'happiness' and 'pleasure' come to the same thing.

until we have tried it. This is Kant's summary of the first-personal aspect of the hedonist thesis:

The property of the subject with reference to which alone this experience can be conducted is the *feeling* of pleasure and displeasure, as a receptivity belonging to inner sense, and thus the concept of what is immediately good would pertain only to that with which the sensation of *gratification* is immediately bound up, and that of what in se evil would have to be referred only to that which immediately excites *pain*.³⁵

Again, Kant is not talking about the impartial variety of hedonism, i.e. classical utilitarianism, whose good lacks an inherent connection with the wills of individual agents. For me, the good would be my actual good, the object that *I* find pleasurable; and what is bad for me would be the object I find painful; *mutatis mutandis*, this holds for you, for me and for everybody else. The good would not only be empirical; it would also be agent relative.

So, what is problematic is not just the contingency of agreement – *if* we reach agreement, and if saying the same thing can *count* as agreement – in matters of pleasure of displeasure. It is contingent that we 'agree', but this is not at the core of Kant's argument against Pistorius, as many interpreters assume.³⁶ The unanimity of inclination falls short of the kind of agreement we might call 'Kantian harmony', which only the formal law of morality can guarantee. If I call something agreeable that just *means* that I take pleasure in it. That is why, appearances notwithstanding, there is no 'peer disagreement' if I say that Kenyan coffee is nice and you say that it is not. We are not contradicting each other at all. Nor are we, strictly speaking, agreeing about anything if we both say that Kenyan coffee is nice. Nothing would be gained even if we had a priori knowledge of the pleasant, i.e. of what pleases you and me respectively.³⁷ We are both merely reporting subjective likes and dislikes, based on the expectation of *our own* pleasure and potentially practical as such.³⁸

35 Kant: KpV, AA 05: 58.17–23.

36 Oliver Sensen argues along these lines with reference to the "unanimity" (*Einhelligkeit*) that is declared contingent in the first chapter of the *Analytic* (KpV, AA 05: 26.12). "Contingent inclinations cannot", Sensen writes, "ground an a priori moral law" (*Kant on Human Dignity*. Berlin/Boston 2011, 25).

37 The conflict will be apparent as soon as we compete for the last cup of coffee left in the pot. In a way our wills would be more likely to be in harmony if our inclinations were directed at *different* things.

38 Kant famously explains the difference between verbal agreement and actual moral harmony with reference to Johann Nikolaus Götz's poem about harmony in marriage, set to music by Joseph Haydn a little later, and even more aptly, in terms of the quarrel between the emperor

8

Kant then proceeds to demolish hedonism as an ethical theory.³⁹ As Lewis White Beck points out, Kant talks “like a twentieth-century philosopher”⁴⁰ in the following passage: Hedonism, he says,

is contrary even to the use of language, which distinguishes the *agreeable* from the *good*, the *disagreeable* from the *evil*, and requires that good and evil always be judged by reason, and hence through concepts, which can be universally imparted [*die sich allgemein mittheilen*⁴¹ lassen], and not through mere sensation, which is restricted to individual subjects and their receptivity [...]. (KpV, AA 05: 58.24–29)

We can rephrase Kant's point in terms of a Moorean ‘open question argument’. There is a question whether what I judge agreeable is actually good. Kant concludes that the good cannot be identical with the agreeable. In some cases we may actually consider pleasure bad, either because it is undeserved or because it is – like *Schadenfreude* – in itself objectionable, and reason can approve of pain, e.g. pangs of conscience. As Beck succinctly puts it: “Reason judges what is desirable; feeling decides what is desired.”⁴²

The Moorean argument against hedonism reemerges in the *Critique of Judgement*. Kant concedes that “in many cases the agreeable and the good

Charles V and his brother Francis I, the king of France, whose will was in perfect agreement: they both wanted Milan (KpV, AA 05: 28.21–22). Even if their desires were less predatory: agreement would at best guarantee action in conformity with duty. It can never be the ground of action that itself is morally good.

39 In what follows, ‘agreeable’ and ‘disagreeable’ etc. should be construed as predicates that – like ‘good’ and ‘bad’ – apply to objects, not to sensations, which are described as pleasurable and painful respectively. So, I find a cup of coffee agreeable iff I enjoy it; I find it disagreeable iff I dislike the taste.

40 Beck, Lewis White: *A Commentary on Kant's Critique of Practical Reason*. Chicago 1960, 132.

41 The original meaning of the German word *mittheilen* – to impart, to give, to share – is barely recognisable today. Like ‘to communicate’, it seems to refer to the process of *verbally* conveying meaning to someone else. But that is not how Kant uses the word. Dieringer has the modern meaning in mind when he asserts that, *contra* Kant, the agreeable and the disagreeable can be ‘communicated’ just like the good and the bad. Dieringer, Volker: *Was erkennt die praktische Vernunft? Zu Kants Begriff des Guten in der “Kritik der praktischen Vernunft”*. In: *Kant-Studien* 93 (2002), 137–157, here: 146. – Others know what I mean when I say that I find rice pudding agreeable. But they may well disagree. And if they do, we will not be able to persuade each other. See also the Canon of the first *Critique*, where Kant says that ‘impartibility’ is the touchstone “whether taking something to be true is conviction or mere persuasion” (KrV, A 820/B 848), and that *knowledge* can be ‘imparted’ or taught (KrV, A 829/B 857).

42 Beck, Lewis White: *A Commentary on Kant's Critique of Practical Reason*. Chicago 1960, 132.

seem to be one and the same”⁴³, particularly if the pleasure in question is sustainable. But, as a matter of fact, it is “merely an erroneous verbal substitution [*fehlerhafte Wortvertauschung*], since the concepts that are properly attached to these expressions can in no way be exchanged for each other”.⁴⁴ Kant adds a novel supplementary argument that once again strengthens his credentials as an analytic philosopher. In the case of the good, he says,

there is always the question whether it is merely mediately good or immediately good (whether it is useful or good in itself), whereas this question cannot even arise in the case of the agreeable, since the word always signifies something that pleases immediately.⁴⁵

Beauty is like the agreeable in that it is also necessarily immediate. But the good is not. So, the good and the agreeable cannot be identical.

9

Does this mean we can finally put the thesis of the priority of the good to rest? Not quite. Kant sets out the one remaining option his opponent might embrace. Pistorius can no longer ascribe goodness to the purposes of human action, to the ends we find agreeable. We may well want to realise the object we expect pleasure from, but the agreeable is not identical with the good, partly because reason cannot establish a necessary connection between specific objects and the pleasure they afford, partly because the use of language indicates that the two are not identical, but perhaps most fundamentally because – unlike the good, which is ‘impartible’ – the agreeable is ineliminably first-personal, subjective and contentious. Still, Kant and Pistorius agree that there is another way in which reason can determine human action, and that there is still room for ‘good’ and ‘bad’, for ‘better’ and ‘worse’.

Actions can be assessed in instrumental terms even if the ultimate end, the agent’s own happiness, cannot. And, in a sense, this involves practical reason.⁴⁶ Instrumental goodness satisfies the criterion of ‘communicability’

⁴³ Kant: KU, AA 05: 207.31–32.

⁴⁴ Kant: KU, AA 05: 207.34–208.1.

⁴⁵ Kant: KU, AA 05: 208.07–10.

⁴⁶ I say ‘in a sense’ because, by the time of the second *Critique*, Kant has come to think that only pure practical reason should actually be called ‘practical reason’. Instrumental practical reason can be reduced to importing the theoretical cognition of causal relations into the human will. See e.g. Kant: KpV, AA 05: 25–26.

or 'impartibility'. You and I can agree about the proper use of means if we disagree about the agreeableness of the end, and even if our very agreement makes us competitors. We can all agree that hot water is good for making coffee,⁴⁷ that cyanide is good for doing in wealthy relatives, and that a compass is good for finding one's bearings. Kant says that

a philosopher who believed that he had to put a feeling of pleasure at the foundation of his practical judging would call *good* what is a *means* to the agreeable, and *evil* what causes disagreeableness and pain; for judging the relation of means to ends certainly belongs to reason.⁴⁸

The goodness in question is the relative, conditional goodness of hypothetical imperatives. Reason plays its role, but it is subservient to inclination.⁴⁹ Given his suggestion that good must always be 'good for something', Pistorius might even be sympathetic.⁵⁰

Unsurprisingly, Kant proceeds to reject instrumentalism as well:

But even though reason alone is capable of coming to see the connection of means with their purposes [...], the practical maxims that would follow from the above concept of the good as merely a means would never contain as the object of the will anything good in itself, but always only something that is good *for something*; the good would always merely be the useful, and that for which it is useful would always have to lie outside the will in sensation.⁵¹

Kant concludes that his opponent is committed to the thesis that there is "nothing at all that is immediately good"; rather, the good is then "to be sought only in the means to something else, namely to agreeableness", which is distinct from goodness.⁵² The conclusion is obviously meant to be

47 I might further be tempted to say that, since I find coffee agreeable, and the preparation of coffee involves heating water in a kettle, that I have a reason to fill my kettle with water, to turn on the gas etc., while you – as you do not fancy a cup of coffee – do not. But as we shall soon see, that is at best misleading and at worst a serious mistake.

48 Kant: KpV, AA 05: 58.31–35.

49 Cf. Kant's account of heteronomy at the end of Section II of the *Groundwork* (GMS, AA 04: 441.03–24).

50 However, it is not clear whether Pistorius would actually embrace instrumentalism now that Kant has removed ends from the realm of goodness. After all, his own brand of instrumentalism depends on the goodness of the end.

51 Kant: KpV, AA 05: 58.36–59.07.

52 Kant: KpV, AA 05: 59.09–11.

absurd. But why should it be absurd to assume that instrumental goodness is the only kind of goodness there is?

10

There are two reasons for us to reject mere instrumentalism on Kantian grounds. First, instrumentalism deprives us of the resources needed to substantiate what Kant regards as our common, everyday conception of morality. As the opening lines of the *Groundwork* illustrate, he takes our regular notion moral goodness to be non-instrumental. So, if instrumental goodness turns out to be the only type of goodness, a conviction central to human morality – that the good will is the one thing that is good unconditionally and absolutely – is unsound because it relies on a conception of value that has been discredited. Ethical theory has to be thoroughly revisionist; or, *contra* Pistorius, we need to put the moral law first.

The appeal of this line of argument is limited, though. The downfall of the good will is bound to upset anyone with only mildly Kantian leanings, but it is unlikely to persuade Pistorius that his criticism of Kant's strategy in *Groundwork* I was unfounded. After all, Pistorius never concealed his scepticism about the notion of an absolutely, non-instrumentally good will. Moreover, it will also be welcomed by the ethical egoist, who can now claim Kantian credentials when he insists that he is being perfectly rational within his own (more or less) ethical framework. Neither Pistorius nor the egoist will readily accept this version of Kant's argument as a *reductio*. The conclusion that 'good' must always be 'good for' is grist to their mill.

Fortunately, however, the idea that goodness is only ever goodness *for something* faces a second, more serious objection. Instrumental goodness is essentially unfounded if it is only ever done for something that is either itself instrumentally justified or – like pleasure – not rationally justifiable at all. This is where the subjective, first-personal nature of pleasure is of critical importance. Consider once again the example of coffee. I can justify my buying beans with reference to the fact that I need to grind them; I can justify my grinding them with reference to the fact that I need ground beans to prepare coffee; I can show that further preparatory steps are rational in much the same way, by adducing objective technical rules. But if you ask me *what purpose* the product – a nice, hot cup of coffee – is meant to serve I can no longer

give you an objective reason. I can only say that I would like to drink it.⁵³ So, if you would like to drink tea and I would like to drink coffee, we can both *instrumentally* justify using water for our cups of tea and coffee respectively. But we cannot justify *taking* the water – if it is, say, a scarce commodity – on impartial grounds, i. e. on grounds we can expect the other person to share.⁵⁴ On the instrumentalist picture, all justification is relative, in terms of what one wants to do. There is no ultimate objective justification.⁵⁵ In practical terms, ‘good for something’ is good for nothing unless it is grounded in the notion of something that is good in itself. Expected pleasure does not qualify.⁵⁶

11

There is another reason why putting a value first is a retrograde step, not progress, for Kant. He conceives of the will as a causal power; and causal powers are determined by laws, not by values. This is clearly stated on the first page of Section III of the *Groundwork*.⁵⁷ If, as a consequence of the will's autonomy, we know that a certain law applies to the human will, our

53 Dieringer also sees the danger of a regress; see: Dieringer, Volker: *Was erkennt die praktische Vernunft? Zu Kants Begriff des Guten in der "Kritik der praktischen Vernunft"*. In: *Kant-Studien* 93 (2002), 137–157, here: 147.

54 Note, however, that impersonal justification (or justifiability) should not be confused with justification (or justifiability) to others. Particularly in the case of duties of virtue, it is perfectly sufficient that I impersonally justify actions to myself. If I fail to help a stranger for no good reason I do not owe him an excuse or an apology. I do not need to justify my action to him (even if it cannot in fact be justified), because he had no right to expect me to help him in the first place.

55 Obviously, it would not help if you were a coffee drinker too, i. e. if we were to agree that coffee is nice. That is not agreement in the requisite sense. We would still be competing for water. I would say ‘I need it!’, you would say ‘But I need it too!’. Even if you need it more badly, that would not as such count as a reason for me. Other than that, we would have nothing to say to each other.

56 In the third *Critique*, Kant similarly emphasises that the practical value of considerations based on sensation is severely limited. What is valued in actions aimed at the agreeable is solely the “gratification they promise”. *How* gratification is achieved does not matter, only *that* the right means are chosen. As a result, human beings could “accuse each other of folly or lack of understanding [*Unverstand*], but never of baseness or malice”. Everyone, “each seeing things his own way”, hurries towards the one goal of individual, subjective gratification (KU, AA 05: 206.12–18). This is precisely the instrumentalist picture of practical reason that is declared unsustainable in the second *Critique*.

57 Cf. Kant: GMS, AA 04: 446.07 and AA 04: 466.15–21.

job is done. By contrast, if we posit a value, we still need to know how this value relates to a law that can determine the will.⁵⁸ Far from solving an old problem, a foundational value creates a new one.

12

Much more needs to be said about the metaethical implications of the primacy of the law; but that must remain the topic for another occasion. For now, let one brief suggestion suffice. While we have done our best to reject a crude realism that seeks to ground ethical norms in a mind-independent conception of value, this does not commit Kant to ‘constructivism’ as it is commonly construed. In fact, we now see why the distinction between realism and constructivism is deeply flawed.

In constructivism, goodness is a matter of *choice*. “For Kant”, Christine Korsgaard famously argues, “what makes the object of your rational choice good is that it is the object of a rational choice”. As human beings, we make choices “and have the attitude that what we choose is good”. In short, “in our private rational choices and in general in our actions we view ourselves as having a value-conferring status in virtue of our rational nature”.⁵⁹ But for Kant, goodness is not about choice.⁶⁰ What constitutes goodness is rational, impersonal judgement, not choice; and – crucially – moral judgement⁶¹ does not determine what we choose to do. Moral failure is voluntary, not cognitive.⁶² Our will has two more or less sovereign masters: reason *and* pleasure, and the good is associated only with reason. As a result, we do

58 This is a version of the argument from the ‘gap’ between the object and the will discussed in § 6 above.

59 Korsgaard, Christine: *Kant’s Formula of Humanity*. In: *Creating the Kingdom of Ends*, Cambridge 1996, 106–132, here: 122–123. – As the phrase “in general in our actions” makes clear, Korsgaard’s thesis is not about rational choice in an elevated, perfectly rational sense. It is about our everyday, conscious and deliberate choices (in which we may well go astray).

60 Obviously, happiness is not good because or when it is chosen. If it were, happiness would almost inevitably be good.

61 Judgement and choice may be more intimately intertwined in the prudential and technical spheres.

62 Those enamoured of -ologies and -isms may still wish to say that Kant is a ‘realist’ about the moral law in that it is not itself optional or chosen; see: Timmermann, Jens: *Emerging Autonomy: Dealing with the Inadequacies of the ‘Canon of Pure Reason’ (1781)*. In: *The Emergence of Autonomy*. Ed. by Stefano Bacin and Oliver Sensen. Cambridge (forthcoming). But it may be best to avoid such labels altogether.

not always act under the guise of the good. Sometimes we act for the sake of pleasures we judge and know to be bad. But that too must remain a topic for another occasion.⁶³

63 I should like to thank audiences at the University of Edinburgh, at the University of Vienna, at Renmin University and at the University of Political Science and Law in Beijing, at the University of Luxembourg and at the University of Halle-Wittenberg for their questions, criticisms and comments on earlier versions of this paper. Particular thanks are due to Violetta Waibel for organising a workshop on Kant's philosophy of value in June 2015 and, of course, the Vienna Kant Congress in September, to Wen Jichang for inviting me to give a series of lectures in Beijing, to Dietmar Heidemann for being my host at the University of Luxembourg with the kind support of the Luxembourg FNR in the autumn of that year, and to Heiner Klemme for his invitation to address the Kant graduate conference in Halle in April 2016 as a keynote speaker. The paper has profited from many Pistorian conversations with Bas Tönissen, Lucas Sierra Velez and Michael Walschots at the University of St Andrews. Finally, I am deeply indebted to my wife, Kate Moran, for countless enjoyable and fruitful discussions of the law and the good.

