Philosophia (2023) 51:609–623
https://doi.org/10.1007/s11406-023-00643-y

Self-Legislation and the Apriority of the Moral Law

Pauline Kleingeld[1]

Received: 1 March 2023 / Revised: 1 March 2023 / Accepted: 3 March 2023 / Published online: 18 April 2023

Abstract
Marcus Willaschek and I have argued against the widespread assumption that Kant claims the Moral Law—the supreme principle of morality—is (or must be regarded as) 'self-legislated'. We argue that Kant instead describes the Moral Law as an *a priori* principle of the will. We also argue that his conception of autonomy concerns not the Moral Law but substantive moral laws such as the law that requires promoting the happiness of others. In the present essay, I respond to the commentary by Alyssa Bernstein and Christoph Hanisch. In response to Bernstein, I clarify the relation between our interpretation of Kant and those proposed by John Rawls and Allen Wood. In response to Hanisch, I argue, among other things, that Kant's defense of the thesis that the Moral Law is an a priori principle of pure practical reason does not mean that practical reasoning and reasons-based action are impossible without it. It means rather that *morally good* and *morally evil* reasons-based action are impossible without it.

Keywords Autonomy · Enabling conditions · Immanuel Kant · John Rawls · Moral Law · Self-legislation

1 Introduction

Kant is widely assumed to have claimed that the Moral Law—that is, the supreme principle of morality—is (or must be regarded as) self-legislated. In our paper 'Autonomy without Paradox: Kant, Self-Legislation, and the Moral Law' (2019), Marcus Willaschek and I argue against this assumption on exegetical and philosophical grounds. Our main thesis is negative, namely that Kant does not claim unequivocally anywhere that the Moral Law is (or must be regarded as) self-legislated and that

✉ Pauline Kleingeld
pauline.kleingeld@rug.nl

[1] Faculty of Philosophy, University of Groningen, Groningen, The Netherlands

his general philosophical commitments do not give him reason to claim this. We also show that Kant instead describes the Moral Law as an a priori principle of the will as such, and we argue that this is insufficiently recognized in the current debate over whether Kant is best interpreted as a realist or a constructivist.

We argue, on exegetical grounds, that Kant's conception of autonomy concerns substantive moral laws, such as the law that requires promoting the happiness of others. Because substantive moral laws are established on the basis of the Moral Law and Kant conceives of the Moral Law as an a priori principle of the will, he conceives of these substantive moral laws as the will's own laws rather than heteronomous impositions. The idea of autonomy thus unites the idea that the will is *subject to* moral laws with the idea that these laws are the will's *own legislation*. In our essay we use capitalization to distinguish the Moral Law (the supreme principle of morality) from substantive moral laws, and I will do the same here.[1]

In this essay I respond to comments from Alyssa Bernstein and Christoph Hanisch. I am grateful to Bernstein for her illuminating discussion of John Rawls, as well as her remarks on our use of the work of John Rawls and Allen Wood, and I thank Christoph Hanisch for developing a stimulating proposal as to how to flesh out the idea that the Categorical Imperative is an a priori principle.

2 Reply to Alyssa Bernstein

Bernstein's critical remarks focus mainly on passages in which we refer to Rawls and Wood. She does not question what Marcus Willaschek and I see as the main thesis of our essay, namely that Kant never claims that the Moral Law is self-legislated. She 'applaud[s] [our] success in providing plausible alternative readings of the relevant passages' (Bernstein 2023, Hanisch 2023) and explains that she does not engage with our discussion of the textual evidence (Bernstein 2023). Furthermore, she does not challenge our interpretation of Kant's conception of autonomy, stating that she agrees with us 'regarding the area of intellectual territory where a satisfactory interpretation of Kant's idea of autonomy is most likely to be developed or discovered' (Bernstein 2023). I highlight these important points of agreement at the outset since Bernstein also claims that the 'main claims' for which we argue 'are false' (Bernstein 2023). In my view, her criticisms concern not so much our main theses as our use of work by Rawls and Wood.

2.1 Rawls and Kant

Bernstein argues that we overlook the fact that Rawls's interpretation of Kant is of the same 'type' as ours. He too reads Kant as saying that the Moral Law is an a priori principle (Bernstein 2023). She argues that Rawls's interpretation of Kant, as presented in his lectures on the history of moral philosophy, thus constitutes a 'counterexample' to our characterization of what we call the standard reading of Kant's conception of autonomy (Bernstein 2023). Furthermore, Bernstein aims to 'defend

[1] We do not do so in quotations, however, for obvious reasons.

John Rawls' from what she sees as our claim that 'Rawls's account of autonomy is paradoxical' and our alleged assertion that 'the moral constructivism attributed by Rawls to Kant is (…) voluntarist' (Bernstein 2023).

In this reply, I would like to clarify the role of our references to Rawls and comment on the relation between Rawls's interpretation of Kant and our own. I would like to point out that we do not claim anywhere that Rawls's account of autonomy is paradoxical, nor do we assert that Rawls sees Kant as a voluntarist. In fact, we do not criticize Rawls at all. We provide two parenthetical references to 'Rawls 1980', that is, the influential essay 'Kantian Constructivism in Moral Theory' (Kleingeld and Willaschek 2019: 1, 14). We also sketch the contours of Rawls's constructivism in a half paragraph, referring to the same essay (Kleingeld and Willaschek 2019: 15) and noting that his version of constructivism is silent on the status of the Moral Law. In what follows, I will start by explaining our aims in including Rawls in our essay and then move to Bernstein's discussion of Rawls's lectures.

The Rawls reference in our introduction follows our observation that Kant's notion of autonomy, as understood on what we call the standard reading, has 'inspired' recent philosophical defenses of Kantian constructivism (Kleingeld and Willaschek 2019: 1). Here we mention, between parentheses, 'Rawls 1980', as well as works by Christine Korsgaard, Onora O'Neill, and Andrews Reath. We mention Rawls's 1980 essay because we believe it has been an influential defense of Kantian constructivism. Our objective in mentioning these authors was to indicate that the relevance of the exegetical issues we discuss in our paper extends beyond Kant scholarship and to anticipate our discussion of constructivism later in the essay.

We were assuming that our later description of Rawls's constructivism would make it sufficiently clear that we do not attribute the standard view to Rawls himself. In this description we state explicitly that his version of constructivism is *silent* on the status of the Moral Law and that it is a theory of how to establish substantive moral obligations (Kleingeld and Willaschek 2019: 15). In this respect we explicitly *contrast* Rawls's position with those held by Korsgaard, O'Neill, and Reath, to whom we do ascribe the view that the Moral Law is self-legislated in some sense.

I am pleased to learn from Bernstein's discussion that in his teaching Rawls presented Kant's theory in a way that seems (at least to an interesting extent) congenial to our interpretation of Kant. In his *Lectures on the History of Moral Philosophy* (LHMP, Rawls 2000) he presents Kant as saying that the Moral Law is an a priori principle of pure practical reason (esp. LHMP 247–252). The Categorical Imperative, he adds, is the procedural representation of the Moral Law as it applies to us (LHMP 247, 251), and this imperative is thus similarly a priori. He states that '[T]he Categorical Imperative is a priori as grounded on pure practical reason' (LHMP, 248) and that it 'is a priori as arising from practical reason' (LHMP, 249). Rawls does not present Kant as arguing that the Categorical Imperative is (or should be regarded as) *self-legislated*.

To the extent that Rawls defended an interpretation of Kant that is in line with ours, Marcus Willaschek and I consider ourselves in excellent company. Rawls offered his students an impressive sketch of the whole of Kant's moral philosophy. It was not his project, however, to argue in any detail against the widespread assumption in the literature that Kant claimed that the Moral Law is (or must be regarded

as) self-legislated. Nor did he accompany his interpretation with discussions of the respective merits of alternative readings proposed in Kant scholarship and Kantian moral theory. This is the project of our essay.

As for Kant's notion of autonomy, here too Rawls's reading agrees with ours, insofar as he sees it as related to substantive moral laws rather than to the supreme principle of morality. In this context, when describing Kant's views, Rawls uses the phrase 'moral law' in the sense of the totality of moral precepts (or 'norms' or 'laws') to which we are legitimately subject, that is, the totality of moral precepts (or norms or laws) resulting from the application of the Categorical Imperative (LHMP 205–206.[2] He writes that we are to regard ourselves as 'legislators, as it were, of the public moral law of a possible realm of ends' (204), that is, as 'legislating law for [our] moral commonwealth' (209), and 'what we legislate, viewed as a moral law for a possible realm of ends, is the whole family of general precepts … that are accepted by the CI-procedure' (205). Thus, when Rawls describes the idea of autonomy as 'the idea of the moral law as a law we give to ourselves as free and equal persons' (204), he is referring not to the supreme principle of morality but to the totality of what Willaschek and I call substantive moral laws. This is further confirmed by the fact that he goes on to describe this idea as 'the thought of our legislating for a possible realm of ends' and the related imperative as demanding that we 'view our maxims as authorized by precepts that could serve as the publicly recognized moral law of such a commonwealth' (204–205).

There are several important differences between Rawls's reading of Kant and our own, however. First, although Rawls considers the Moral Law to be a priori, he does not attribute much importance to this aspect of Kant's theory. Rawls hardly discusses how the claim is to be understood, and this turns out to be deliberate. He explains that he regards the claim as being of only peripheral importance:

> [I]n presenting Kant's moral philosophy, I have played down the role of the a priori and the formal, and I have given what some may think a flat reading of the categorical imperative as a synthetic a priori practical proposition. These things I have done because I believe that the downplayed elements are not at the heart of his doctrine. (LHMP 275)

In Rawls's eyes, the procedure Kant offers for establishing substantive moral principles, as developed through the formulas of the Categorical Imperative, is of much greater importance. This procedure inspired Rawls's own constructivist theory.

On the view that Marcus Willaschek and I defend, by contrast, Kant's argument that the Moral Law is an a priori principle of pure practical reason is undeniably crucial to his doctrine. Kant's philosophical ambition was to develop a metaphysics of

[2] Rawls uses 'moral law' in the singular in at least three senses. He uses it to refer to (1) the 'supreme principle of morality' (e.g., LHMP 147), to which Willaschek and I refer as the Moral Law (capitalized); (2) the totality of moral 'precepts' or 'norms' or 'laws' to which we are legitimately subject (e.g., LHMP 205–206, analogous to 'the law of the state' as the totality of specific state laws), that is, the totality of substantive moral laws (or 'precepts' or 'norms') resulting from the application of the Categorical Imperative; and (3) one element of this totality, that is, one of the moral laws of the ideal moral commonwealth (e.g., LHMP 208).

morals. To achieve this goal, he first needed to 'search for and establish' the supreme principle of morality (GMS 4:392),[3] which we refer to as the Moral Law. Only on this basis would Kant subsequently be able to construct a systematically organized body of substantive moral laws—or at least this was his ambition (Kleingeld and Willaschek 2019: 13–14). Rawls was not in the business of developing a metaphysical system, and hence the a priori status of the Moral Law is not at the center of his interest in Kant.

Another difference between Rawls's interpretation of Kant and our own is the fact that in his discussion of autonomy he at times seems to divorce the role of 'legislator' from that of 'subject'. On our view, by contrast, Kant's notion of autonomy serves to make it possible to regard the will as *both* subject and legislator of moral laws, such that moral laws can be regarded as the will's own legislation. Rawls does not see it that way, as indicated by his introduction of the Formula of Autonomy:

> With the autonomy formulation, there is a shift of point of view (…): we now come back to viewing ourselves not as subjects of the moral law, but as legislators, as it were, of the public moral law of a possible realm of ends. (LHMP 204)

Rawls repeatedly makes similar comments elsewhere in the lectures (e.g., LHMP 183). On the account that Willaschek and I develop, by contrast, Kant's conception of autonomy is meant precisely to bring together the role of the will as subject and its role as legislator: in viewing the will as both subject to and legislator of moral laws, we view the will as subject to its own legislation, that is, as having autonomy, rather than as being subject to alien dictates.[4]

Despite these differences between Rawls's presentation of Kant and our own, I am pleased to see that Rawls acknowledges the a priori character of the Moral Law (or, as he most often writes, Categorical Imperative) and does not (as far as I have been able to ascertain) read Kant as saying that the Moral Law is or must be regarded as self-legislated. In contrast to Bernstein, however, I do not believe that our characterization of the 'standard reading' in the Kant literature is undercut by one counterexample found in lectures to students, even when that example is Rawls. A plausible criterion by which to judge whether a particular interpretation is a standard reading is whether authors assert it without arguing for it, that is, whether they take it for granted and assume that their readers will do so too. Willaschek and I believe this is the case with the assumption that Kant claimed that the Moral Law is (or must be

[3] All references to Kant's work cite volume and page number of the *Akademie-Ausgabe*. Translations are taken from the edition listed in the References (Kant 1996), though I have sometimes altered them. Abbreviations: GMS: Groundwork for the Metaphysics of Morals (*Grundlegung zur Metaphysik der Sitten*); KpV: Critique of Practical Reason (*Kritik der praktischen Vernunft*); RGV: Religion within the Boundaries of Mere Reason (*Die Religion innerhalb der Grenzen der bloßen Vernunft*); V-NR/Feyerabend: Feyerabend Lectures on Natural Law (*Naturrecht Feyerabend*).

[4] Elsewhere I argue that this conception of autonomy disappears from Kant's later writings. In the *Metaphysics of Morals*, where Kant distinguishes *Wille* and *Willkür*, he no longer speaks of autonomy as a property of the will (*Wille*), as he had done in the *Groundwork*. For discussion, see Kleingeld (2018) and 'Kant's Formula of Autonomy: Continuity or Discontinuity?' in this special issue (Kleingeld 2023).

regarded as) self-legislated. As far as we could see when we wrote our essay, this assumption was not a matter of scholarly debate in the Kant literature.

In Sect. 8 of Bernstein's comments, it becomes clear why she believes that we criticize Rawls's interpretation of Kant. She writes that our phrase 'John Rawls's Kantian constructivism' is 'ambiguous between (a) Rawls's interpretation of Kant's moral philosophy and (b) Rawls's own political philosophy' and takes us to mean the former (Bernstein 2023). However, by the phrase 'John Rawls's Kantian constructivism' we mean Rawls's own constructivism. Thus, in response to Bernstein's complaint that we do not make clear what problems we see with Rawls's interpretation of Kant, I can only respond by saying that we do not discuss or criticize his interpretation of Kant at all in our essay. Similarly, when she writes that it 'seems reasonable to infer' that we 'implicitly ascribe to Rawls' (Bernstein 2023) the view that Kant defends a form of voluntarism, again I can only reply that we neither discuss nor criticize Rawls's interpretation of Kant. We merely present, rather briefly, Rawls's constructivist theory.[5]

Bernstein further argues that Willaschek and I conflate 'questions about moral obligation' with 'questions about moral law' and that answering these questions requires more argument than we offer. For example, explaining moral obligation would require a discussion of moral psychology and freedom of the will, neither of which we provide. In reply, I would like to point out that it is not our aim in the essay to provide a complete interpretation of Kant's arguments in Parts 2 and 3 of the *Groundwork*. Our aim is restricted to showing that Kant describes the supreme principle of morality not as self-legislated but as a priori, and that he describes autonomy as relating not to the Moral Law but to substantive moral laws. When we speak of the 'normative force' or 'obligatory force' of the Moral Law, we do so in the context of this narrower project. Here we ask whether Kant describes the supreme principle of morality as deriving its normative force from an act of self-legislation. A full treatment of Kant's account of moral obligation (including a discussion of moral psychology and freedom of the will) is not necessary for answering this question.

2.2 Wood and Kant

We introduce Allen Wood as one of the authors who interpret Kant as a metaethical realist. Our main point is to show that Wood, in line with the standard reading, interprets Kant's talk of self-legislation as relating to the Moral Law. We also mention that he (rightly) observes that Kant often uses terms such as 'considering' or 'regarding as' in this context. On Wood's view, Kant holds that the Moral Law must be *regarded as* self-legislated, not that it *is* self-legislated (Wood, 2008: 116–117). According to Wood, this makes it possible to reconcile Kant's metaethical realism with his talk of self-legislation in relation to the Moral Law and with his assumption that the Moral Law is a natural law ('natural' here in the sense of the natural law tradition) (Wood, 2008: 108–109). Our disagreement with Wood concerns his presupposition that

[5] Also, we do not claim anywhere that 'all scholars (…) who interpret Kant as a moral constructivist ascribe to him the paradox of legislation' (Bernstein 2023).

Kant's talk of self-legislation *relates to the Moral Law*; on our view, in the passages Wood refers to, it relates to substantive moral laws.

Bernstein does not criticize our discussion of Wood's interpretive proposal. Rather, she focuses on our quoting Wood as stating that there is a 'serious tension in the Kantian idea of autonomy' (Wood, 2008: 106, quoted by Bernstein 2023). We quote Wood's statement to support our contention that it is difficult to give coherent sense to the idea of autonomy on the standard reading. Bernstein argues that Wood is here referring not to *Kant's* idea but only to the (later) *Kantian* idea of autonomy (Bernstein 2023) and hence that we misrepresent Wood's position. It is not clear to me that this is right, however, as I will explain.

Willaschek and I quote from the first page of Wood's chapter on autonomy. He starts this chapter with two Kant quotations and an outline of what he sees as a tension in the 'Kantian idea' of autonomy, a tension between the Moral Law's objective validity and its being self-legislated. He then notes that the term 'autonomy' was later separated from 'the Kantian doctrines that were its original home'. Wood's use of the term 'Kantian' here is certainly ambiguous. Given the quotations from Kant and the reference to the 'original home' of the idea, however, it seems natural to read Wood as referring to *Kant's* idea (which had its original home in *Kant's* doctrines, from which later theorists removed it). This is also confirmed by Wood's subsequent discussion. After outlining the 'tension' or 'dilemma' *in the idea* of autonomy and discussing how later authors transformed the idea, Wood transitions to an exegetical discussion of Kant's own conception of autonomy by pointing to what he calls Kant's 'own choice' (Wood, 2008: 110) for dealing with the alleged tension. Kant's choice, Wood argues, is to take the idea of self-legislation of the Moral Law non-literally (as mentioned above). Thus, our presentation of Wood does not seem to me to be a misrepresentation: Wood does assert that there is a serious tension *in the idea of autonomy* as introduced by Kant, even as he argues that Kant chose to resolve it by taking one half of the idea non-literally.

In conclusion, I do not believe that Bernstein shows, as she claims, that the main theses for which we argue are false. Our main theses are that Kant does not claim that the Moral Law is (or must be regarded as) self-legislated and that his terminology of the will's 'self-legislating' instead concerns substantive moral laws. Bernstein does not dispute these theses. She questions our use of Rawls's and Wood's work, and I hope to have provided sufficient clarification in response.

3 Reply to Christoph Hanisch

Hanisch points out that Bernstein, Willaschek, and I show that the differences between realism and constructivism are less pronounced than is often assumed. I agree with his assessment, and I start my reply by developing this point further. I then turn to Hanisch's proposal that we should conceive of the Categorical Imperative as an enabling condition for action and action as an enabling condition for the Categorical Imperative. Finally, I address his position on the issue of whether the Categorical Imperative, on Kant's account, is self-legislated.

3.1 The Apriority of the Moral Law as Common Ground

Hanisch starts with the insightful remark that Rawls and Wood agree on the metaethical status of the Moral Law to a much higher degree than is commonly assumed (including by Wood himself). I would like to extend the scope of this remark by showing, on the basis of the work of Andrews Reath, that it also applies to at least some authors who defend a constitutivist reading of Kant.[6]

In his lectures, Rawls describes the Moral Law as an a priori principle that arises from the nature of reason. Wood, who could not be clearer that he interprets Kant as a metaethical realist (Wood, 2008: 112–113), similarly asserts that the Moral Law is an a priori principle that arises from the nature of reason. Wood writes that on Kant's account practical reason is a power or faculty and that faculties have constitutive norms that guide their proper functioning. The highest norm of the faculty of practical reason is the Moral Law (Wood, 2008: 111–116, esp. 115–116; note that the terminology of constitutive norms is Wood's own). Andrews Reath, who interprets Kant as a metaethical constitutivist (most recently Reath, forthcoming), similarly holds that the fundamental principle of morality is a constitutive norm of practical reason and arises from the nature of reason (Reath, 2006: 92–120; Reath 2013; Reath forthcoming).

Wood and Reath disagree on many other interpretive issues, such as the status of the Formula of Universal Law. According to Wood, the Formula of Universal Law is an incomplete version of the supreme principle of morality, as it is only the 'formal' part of the foundation of Kant's moral theory and needs to be supplemented with 'matter' furnished by the value of humanity. Wood therefore claims that Kant's moral philosophy is grounded on values (Wood, 2005, 129). Reath disagrees and holds that the Formula of Universal Law is itself the supreme principle, not merely an incomplete version of it. Nevertheless, they both hold that the supreme principle of morality is an a priori principle arising from the nature of practical reason.

Wood and Reath have different ways of accounting for the (alleged) status of the Moral Law as 'self-legislated'. Both relate the idea of self-legislation in Kant to the *Moral Law*.[7] Both argue forcefully against voluntaristic interpretations of the idea. Both also argue that the Moral Law is *not literally* self-legislated, although they do so in somewhat different ways. Wood understands Kant as using the term 'legislating' (in 'self-legislating') in the ordinary sense of enacting a law that was not previously in force, thereby creating an obligation (Wood, 2008: 112–113). He takes Kant to argue that the Moral Law is not literally self-legislated in this sense but must be '*considered*' *or* '*regarded as*' *self-legislated* (as mentioned above). Reath points to these locutions in Kant's texts as well, but he also offers an alternative understanding of the *term* 'legislating' itself. He argues that the Categorical Imperative is self-legislated in the sense that it 'is the law that emerges from the very nature of rational volition'

[6] The Categorical Imperative is the prescriptive expression of the Moral Law. With respect to human beings, both terms can be used. Willaschek and I mainly refer to the Moral Law; Hanisch mainly refers to the Categorical Imperative; some of the other authors discussed use the former, others the latter. In my reply, I will therefore use both, without always explicitly distinguishing between them.

[7] I here bracket the issue of whether they also relate it to substantive moral laws; Reath explicitly relates it to both.

(Reath, 2006: 99; cf. Reath 2018: 182–183). He further argues that if an agent's reasons for complying with a law are the same reasons that would lead a legislator to enact it, then the agent subject to the law 'goes through the same deliberative process as a legislator does in enacting it', and thus 'the subject may be regarded as legislating' (Reath, 2006: 103). None of Reath's descriptions involve 'self-legislation' of the Moral Law in the literal sense, however. Thus, while there are certainly differences between Wood's and Reath's interpretations of the (alleged) non-literal 'self-legislation' of the Moral Law, these differences are subtle and not nearly as substantial as they are sometimes taken to be.

More importantly, if the exegetical argument that Willaschek and I offer is convincing and Kant's terminology of self-legislation does not concern the Moral Law (the supreme principle of morality) in the first place, then any disagreement concerning Kant's (alleged) description of the Moral Law as self-legislated is moot, since Kant does not describe the supreme principle of morality as self-legislated at all. What is left, then, is the claim that Kant sees the Moral Law as an a priori principle arising from the nature of pure practical reason, and *this* is a claim with which Rawls, Wood, Reath, Willaschek and I all agree.

In other words, if the surface disagreement concerning the modality of the alleged self-legislation of the Moral Law is removed, then we find common ground in the claim that Kant conceives of the Moral Law as an a priori principle of pure practical reason. There remain plenty of other issues on which those on either side of the current 'realist' and 'constructivist' divide will continue to disagree. But I agree with Hanisch that it could be fruitful for the debate if more attention were focused on this important point of consensus.

3.2 The Categorical Imperative as an Enabling Condition for Action

The body of Hanisch's comments consists of a constructive twofold proposal concerning how best to characterize the role of this a priori principle in the agent's practical reasoning and action. He proposes that we conceive of the Categorical Imperative as an enabling condition for reasons-based action and reasons-based action as an enabling condition for the Categorical Imperative.

Hanisch borrows the notion of an enabling condition from Jonathan Dancy to refer to conditions that are 'not considered an essential part of the explanation of an event' but necessary for the event to occur (Hanisch 000; cf. Dancy 2000: 127–130). Hanisch gives the example of the presence of oxygen when an unattended toddler is playing with matches and accidentally sets a house on fire. Although the presence of oxygen makes it possible for the fire to occur, under ordinary circumstances its presence would not be considered an essential part of the explanation of what caused the house to burn.

As Hanisch explains, Dancy introduces this notion to make a point about the role of beliefs in the explanation of actions. On Dancy's view, beliefs are enabling conditions for actions, not explanatory reasons; when we ask for an agent's reasons for action, we are asking for (believed) *facts* (Hanisch 2023). You ask me: Why did you go to that far-away store? And I explain: Because it has the best fruit. What explains my action is *the fact that it has the best fruit*. My having the *belief* that it does is

an enabling condition for my going there, since I would not go there if I did not believe they had the best fruit. But I did not go to the far-away store for the reason *that I have this belief*. Or so Dancy argues.[8] Hanisch's point is not to defend Dancy's theory, however, but merely to introduce the notion of an enabling condition, which he employs for his own purposes.

The first component of Hanisch's proposal is that we should understand the Categorical Imperative as an enabling condition for reasons-based action. On his understanding, the Categorical Imperative functions as a general coherence constraint that enables practical reasoning in the first place. He formulates this as follows:

> The presence (…) of the Categorical Imperative, understood as a structural feature of the first personal deliberative stance of every agent, is necessary for the agent to come up with reasons in the course of confronting the task of deciding what to do and how to act. Reasons for action, so understood, come with the minimal characteristics of internal coherence and consistency, i.e., the Imperative's demand for law-likeness in one's practical principles. (…) [W]hile the deployment of the Categorical Imperative is necessary for practical reasons to be reasons, it is not an explicit and, thus, foregrounded principle at this point. (Hanisch 2023)

Hanisch also offers a 'genealogical' account of how the Categorical Imperative becomes foregrounded in practical deliberation and applied in practice, as I discuss below.

The second component of Hanisch's proposal is the claim that phenomenal actions are enabling conditions for the Categorical Imperative, by which he means that empirical actions, in space and time, are its 'existence condition' (Hanisch 2023). He expects that this part of his proposal will be met with 'understandable skepticism' by Kantians, and I do indeed see how it could lead to worries about its compatibility with Kant's transcendental idealism.

Hanisch's aim, however, is not so much to *argue for* this twofold proposal as to *use* it to sketch the contours of an account of how the Categorical Imperative might function as a normative a priori principle of pure practical reason. He does not have the space, of course, to provide the complete argument for his proposal in the context of brief comments. Therefore, I will focus not so much on the argument in support of his proposal as on some of its salient features, by comparing it to what I see as Kant's view.

First, whereas Hanisch's account of the Categorical Imperative as an enabling condition makes it impossible for it to play the role of an explanatory reason, it seems to me that Kant does make room for such a role. To be sure, the Categorical Imperative does not play *any* role in the explanation of an action when 'explanation' is taken in Kant's strict sense of an account of the necessity of a phenomenon in terms of laws

[8] A full explanation of why I go to the far-away store should of course mention additional elements of my practical reasoning, such as the fact that produce quality is more important to me than saving time shopping, and it should include additional enabling conditions such as my having money to buy groceries, time to go to the far-away store, and the mobility required for shopping. I do not pursue this any further here, let alone address the substance of Dancy's theory.

of nature. After all, the Categorical Imperative is not a law of nature. When it comes to explaining phenomenal actions as such, it cannot even play the role of an enabling condition, let alone the role of an explanatory reason. But Hanisch seems to use the notion of an 'explanatory reason' not in Kant's strict sense of 'explanation' but in the sense of an account of the practical reasoning that underlies a given action – an account that need not have the epistemic status of knowledge. If we use the notion in this sense, it seems to me that Kant permits reference to the Categorical Imperative as an explanatory reason, at least in some cases, contrary to Hanisch's proposal.

When we assume, as Kant believes we sometimes do, that someone performed an action from duty – that is, performed it because the Categorical Imperative demanded it – then we take the Categorical Imperative (or more precisely the agent's recognition of its authority) to be an 'essential part' of the account, not merely an enabling condition. Consider a person who does not feel like helping others in need (cf. GMS 4:423) but who recognizes both that the Categorical Imperative demands that he help them and that this demand has overriding normative force. If and when this person decides to help for this reason, he may indeed believe he does so from duty, that is, for the reason that the Categorical Imperative demands it. Here the agent's recognition of the authority of the Categorical Imperative is an 'essential part' of his own account of the action in terms of his underlying practical reasoning. If the Categorical Imperative is merely an enabling condition, however, as Hanisch proposes, then it cannot play this role. Thus, it seems that Kant does not conceive of the Categorical Imperative as merely an enabling condition.

Second, Hanisch's proposal seems to imply that only *morally good reasons* are reasons and that only *morally good action* is action. If so, then his proposal differs from Kant's position in this regard as well. Hanisch understands the 'deployment' of the Categorical Imperative as 'necessary for practical reasons to be reasons' (quoted above). One important question here is whether by the 'deployment' of the Categorical Imperative as a normative 'structural feature of the first personal deliberative stance' Hanisch means the agent's *recognition of its authority* or rather their *compliance with its demands*. His wording suggests the latter (see the block quote above). He writes that reasons, in order to be reasons at all, must display internal coherence and consistency, and he sees this as what the Categorical Imperative demands. He further calls the Categorical Imperative a normative background constraint on practical reasoning that fulfills a role comparable to the role of the laws of logic in thought (Hanisch 000). All of this means that, on Hanisch's view, agents must reason in conformity with the demands of the Categorical Imperative if they are to reason at all. An agent who decides to act from self-interest rather than from duty would not really be engaging in genuine practical reasoning and genuine action.

Kant, by contrast, insists that humans can *reason* and *act* on the basis of morally impermissible action principles. In the *Groundwork*, he famously distinguishes 'actions that are contrary to duty' (*pflichtwidrige Handlungen*) from actions that are 'in accord with duty' and actions that are done 'from duty' (GMS 4:397). All three are *actions.* In the equally famous illustration, the prudent merchant who acts in accord with duty (but not from duty) *reasons* from the maxim of self-interest (GMS 4:397). Maxims are the major premises of an agent's practical reasoning. What Kant calls an evil action (*böse Handlung*, e.g., KpV 5:153, RGV 6:39, 41) is an action performed

on the basis of an evil maxim (*böse Maxime*, RGV 6:20), that is, a maxim that is morally impermissible. Kant assumes that we encounter many actions in everyday life that lead us to attribute such impermissible maxims to their agents. He denies, of course, that we can ever know with certainty the reasoning that underlies a particular observable action. Our assumptions about the motives of others (and our own) have an epistemic status that is much weaker than that of knowledge. But, as I believe the textual evidence provided here shows, his philosophical theory clearly allows for the possibility that agents reason and act on the basis of morally impermissible maxims.

Furthermore, Kant claims that reasons-based action is possible *without* the Categorical Imperative. At the time of writing the *Groundwork*, he asserted that if human beings did not have freedom, they could do on the basis of reason (*durch die Vernunft*) what animals do on the basis of instinct (V-NR/Feyerabend 23:1322). Humans would have reason, but they would not have a will of their own; they would follow the will of nature (V-NR/Feyerabend 23:1322). Since acting from respect for the Categorical Imperative requires freedom, Kant's claim implies that without the Categorical Imperative it would be possible for action to be *reasons-based* but not *morally good or evil*.

In a similar vein, in the *Religion* Kant distinguishes between the '*predisposition for humanity*' and the '*predisposition for* [moral] *personality*' (RGV 6:26–28), explaining the former by saying that it is the predisposition for reasoning instrumentally (RGV 6:28). This again indicates that Kant held that practical reasoning is possible independently of the Categorical Imperative. In a famous footnote Kant further explains this by saying that, conceptually speaking, it does not follow from the fact that a being has reason that it can determine its will by means of anything other than inclination-based incentives. The most reasonable being in the world could in principle lack all awareness of (even the mere possibility of) the Moral Law and could still act on the basis of reasons (RGV 6:26, n.).

Thus, Kant does not seem to conceive of the Categorical Imperative (or Moral Law) as an enabling condition for reasons-based action as such. Kant's defense of the thesis that the Moral Law is an a priori principle of pure practical reason does not entail that practical reasoning and reasons-based action as such would be impossible without it. It only means that *morally good and morally bad (evil)* reasoning and action would be impossible without it.

Hanisch presents his proposal not as a matter of Kant exegesis but as a contribution to contemporary Kantian action theory, and my comparison with Kant's account in this section is not meant as a refutation of his proposal. Yet his proposal is further removed from Kant's account of action in general, and of moral action in particular, than it may seem at first. I believe the same point can be raised vis-à-vis the core claims defended by leading Kantian constitutivists who similarly see the Moral Law as a necessary condition for practical reasoning and action, but I defer a discussion of this issue to another occasion.

3.3 Hanisch on self-legislating the Moral Law

Hanisch includes an account of how deliberating agents become reflectively aware of the Categorical Imperative. He writes that the Categorical Imperative, which on

his proposal is initially a background constraint, becomes foregrounded, acknowledged, and affirmed (or reaffirmed) by agents when they 'conceptualize the "fact" of encountering the normative force of the Categorical Imperative' (Hanisch 2023). He writes that acknowledging and affirming (or reaffirming) the normative force of the Categorical Imperative 'can be called an act of "self-legislating" the Categorical Imperative' (Hanisch 2023). On his view, once the Categorical Imperative is self-legislated in this way, it becomes possible to derive substantive duties on its basis.[9]

I would first like to note that if the Categorical Imperative were to play an enabling role in an agent's practical reasoning, this would *already* require the reasoner's recognition of its authority; otherwise, the imperative could be regarded as mere noise. This is to say that the mere *presence* of the Categorical Imperative (and even its mere presence as a 'structural feature', as Hanisch puts it in the block quote above) does not suffice for it to serve as an enabling condition for practical reasoning. For the Categorical Imperative to fulfill this role in practical reasoning, the reasoner must at least implicitly *already recognize* it *as binding*, that is, as a norm that ought to guide their practical reasoning.

On Hanisch's account, however, this recognition of the authority of the Categorical Imperative is introduced only as the *second* part of his normative genealogy, and he describes this as an act of 'self-legislating' the Categorical Imperative. Hanisch does not take this to mean *enacting a law* but rather, as the quotations indicate, *acknowledging the normative force of an existing law* (here the Moral Law or, in its application to finite rational beings, the Categorical Imperative). He emphasizes that self-legislating, thus understood, is not problematically arbitrary and voluntaristic, and he expresses agreement with Willaschek and me on this issue. We do indeed emphasize that consciousness of moral obligation involves *recognition* of the normative force of the Moral Law, but we would insist that this recognition must be present from the start (as mentioned above). More importantly, in our essay we argue *against* the common assumption that Kant describes this recognition *in terms of self-legislation of the Categorical Imperative* (Moral Law, supreme principle of morality).

Hanisch refers to only one passage in support of his interpretation that Kant describes the Categorical Imperative as 'self-legislated'. In this passage, Kant refers to '*dem moralischen Gesetz*' and writes that a person who fulfills his duties is not only subject to it but also 'lawgiving with respect to it'. Hanisch's reason for thinking that '*dem moralischen Gesetz*' refers to the Moral Law (the supreme principle of morality) rather than a moral law (a substantive moral principle) is that Kant uses the singular. In our essay, however, Willaschek and I argue that Kant sometimes uses 'moral law' as a *generic singular* to refer to substantive moral laws in general (Kleingeld and Willaschek 2019: 5–6). The generic singular sense of 'law' is the sense in which a police officer might say that you must 'obey the law' (in the singular) even if you have broken multiple traffic laws by speeding through a red light while intoxicated and without a driver's license. We also indicate why we believe that the singular, in the passage to which Hanisch refers, is best read as a generic singular (2019: 8–9).

[9] Hanisch is careful to note that this is not an empirical psychological development, let alone a temporal sequence (Hanisch 2023, note 18).

Thus, the mere fact that Kant uses the singular here does not seem to be sufficient to show that Kant indeed claimed that *the Moral Law* (or *Categorical Imperative*, as Hanisch prefers) is 'self-legislated', or that it 'can be called' or 'should be regarded' as such. With that said, we welcome further discussion of these and related interpretive issues to assess the pros and cons of different readings. Our aim in the essay is not to show that other readings of Kant are impossible but to articulate an interpretation that—as we believe—has exegetical and philosophical advantages and has been overlooked thus far.

I am therefore immensely pleased that Bernstein and Hanisch agreed to comment on our essay and am grateful for their comments. I hope this discussion will continue.

Acknowledgements I thank Sorin Baiasu and Micha Gläser for editing this special issue. I am grateful for helpful comments from Marcus Willaschek, Micha Gläser, Sorin Baiasu, Leon van Rijsbergen, and Janis Schaab. I thank the Dutch Research Council (NWO) for financial support.

Competing Interest The author declares that she has no conflict of interest.

References

Bernstein, A. R. (2023) 'Autonomy and Objective Moral Constructivism: Rawls versus Kleingeld & Willaschek'. *Philosophia*. https://doi.org/10.1007/s11406-023-00620-5.

Dancy, J. (2000). *Practical Reality*. Oxford: Oxford University Press.

Hanisch, C. (2023). 'The Categorical Imperative *in* Action:Enabler and Enablee of Self-Legislation'. *Philosophia*. https://doi.org/10.1007/s11406-023-00637-w.

Kant, I. (1996). *Practical Philosophy*. Translated and edited by M. J. Gregor. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511813306.

Kleingeld, P. (2018). 'The Principle of Autonomy in Kant's Moral Theory: Its Rise and Fall'. In E. Watkins (Ed.), *Kant, Persons, and Agency* (pp. 61–79). Cambridge: Cambridge University Press.

Kleingeld, P. (2023). 'Kant's Formula of Autonomy: Continuity or Discontinuity?' *Philosophia.* https://doi.org/10.1007/s11406-023-00642-z.

Kleingeld, P., & Willaschek, M. (2019). 'Autonomy without Paradox: Kant, Self-Legislation and the Moral Law'. *Philosophers' Imprint*, *19:6*, 1–18.

Rawls, J. (1980). 'Kantian Constructivism in Moral Theory'. *Journal of Philosophy*, *77*, 515–572.

Rawls, J. (2000). *Lectures on the History of Moral Philosophy.* Edited by B. Herman. Cambridge, MA: Harvard University Press.

Reath, A. (2006). *Agency and Autonomy in Kant's Moral Theory*. Oxford: Oxford University Press.

Reath, A. (2013). 'Kant's Conception of Autonomy of the Will'. In O. Sensen (Ed.), *Kant on Moral Autonomy* (pp. 32–52). Cambridge: Cambridge University Press.

Reath, A. (Forthcoming). 'Kant's Moral Constitutivism'. In Bacin, S., & Bagnoli, C. (Eds.), *Reason, Agency, and Ethics: Kantian Themes in Contemporary Debates*. Oxford: Oxford University Press.

Wood, A. W. (2005). *Kant*. Malden: Blackwell.

Wood, A. W. (2008). *Kantian Ethics*. Cambridge: Cambridge University Press.