

Autonomy and Moral Rationalism: Kant's Criticisms of 'Rationalist' Moral Principles (1762-1785)

Stefano Bacin

in Stefano Bacin, Oliver Sensen (eds.), *The Emergence of Autonomy in Kant's Moral Philosophy*, Cambridge, Cambridge University Press, 2019, 48-66

— Please cite published version —

1. *Kant's Taxonomy of the Principles of Morality*

The idea of autonomy of the will emerges in Kant's moral philosophy as the solution to a crucial issue that Kant considers unsolved by previous moral philosophers: the question of the status and the foundations of moral obligation (cf. 2:298). The account first presented in the *Groundwork* must be interpreted with regard to Kant's diagnosis of the state of the discussion to which he aimed to contribute. His examination of the previous views does not provide a mere background, but instead represents a momentous aspect of the development of Kant's own account. A close connection between Kant's thesis of the autonomy of the will and the assessment of the previous accounts of morality is clearly suggested both in the *Groundwork* and the second *Critique*, where critical remarks on previous views immediately follow the presentation of the idea of autonomy. Kant's notion of autonomy provides a new vantage point from which the shortcomings of the previous views are better seen. Kant accordingly rejects all previous views on the charge of heteronomy (cf. GMS 4:441, KpV 5:33, 29:629).

Kant's remarks in the *Groundwork* and the second *Critique* on previous moral principles derive from the thorough appraisal that can be followed in earlier unpublished texts. Many of Kant's private notes and, more extensively, each version of his lectures on ethics discuss the relevant viewpoints on the foundations of morals. While the terminological contrast between autonomy and heteronomy is still absent from the notes and the lectures prior to

the *Groundwork* and emerges only later, in the Mrongovius II notes (cf. 29:629),¹ Kant's criticisms prior to 1785, show how that contrast emerged. Before Kant put the previous views together as heteronomous views, he addressed them with specific criticisms, which, unlike the all-inclusive heteronomy objection, reveal specific weaknesses of each view and, more importantly, how those views include insights that proved helpful in the development of Kant's mature conception. His discussion of the previous views is extraordinarily valuable for understanding the idea of autonomy not only because it sheds light on his desiderata for a solution to the central issue of moral philosophy, but also because it clarifies how Kant's idea of autonomy relates to earlier views. As I shall argue, Kant's examination of rationalist views directly contributes to the development of his view of the autonomy of the will.

Before addressing this issue, however, the general character of Kant's critical remarks on the previous views must be briefly explained. In the lectures, Kant's examination of the previous accounts mainly occurs in the section "On the Supreme Principle of Morality", and, occasionally, in the section "On Law".² Remarkably, those discussions belong to the very few parts of the lectures that do not directly correspond to Baumgarten's textbooks Kant comments on. The whole section "On the Supreme Principle of Morality", which *prima facie* appears an unexpected duplication of the previous section "On the Principle of Morality", is an addition of Kant's to Baumgarten's order, which Kant closely follows in the rest of the lectures.³ The insertion of a specific section in the schedule of the lectures, however, must not have been primarily motivated by the aim of supplementing the textbook with a historical summary.⁴ While Kant's logic and metaphysics lectures have sections dedicated to fairly

¹ On the special significance of the Mrongovius II notes see Timmermann 2015.

² Cf. 27:9f; 27:135-7; 27:274-8; Kaehler 55-73; 29:620f. I shall mainly focus on the lecture notes up to 1785, relevant to the development of the idea of autonomy. The later Vigilantius notes briefly rehearse the same arguments as the earlier lectures (27:497-9).

³ See Reich 2001, 387. As Reich observes, the title "Of the supreme principle of morality" anticipates Kant's phrasing of the aim of the *Groundwork*: "the identification and corroboration of the supreme principle of morality" (4:392). On the relation between Kant's ethics lectures to Baumgarten's works see Bacin 2015.

⁴ Wood 2015: 124, stresses that Kant's remarks on previous moral theories sketch a brief "history of ethics", though "not in the narrative-chronological sense, but in the taxonomical sense".

comprehensive historical sketches, the remarks on previous moral theories are particularly selective. Kant only mentions a few selected views, often without even naming their advocates, and does not present them as stages of a continuous development. Instead, his critical remarks concentrate on the modern views, which propose solutions to the question of moral obligation. While Kant's comments did provide his students with some background on the history of the discipline, their main aim was to map the relevant conceptions of the principle of morality. This purpose explains why analogous remarks on heteronomous theories occur in the *Groundwork* and the second *Critique*, where a genuinely historical summary of the conceptions of morality would have been out of place.⁵ Thus, the evidence suggests that the discussion of the previous principles is directly significant for how Kant develops his own position.

Kant organizes his examination according to a taxonomy that, remarkably, did not change significantly through the decades. Kant first separates ancient views, which he takes to be focused on determining the contents of virtue and the highest good, and then modern views, which he presents as primarily concerned with the principle of morality, or what he calls the "sources and principles of moral evaluation" (27:106; cf. e.g. Refl. 6760, 19:151).⁶ Focusing on modern views, he then distinguishes between empirical and rationalist views: the former propose principles that can only be cognized empirically, while the latter put forward principles that require rational cognition. This distinction may appear in contrast with the heteronomy objection presented in the works of the 1780s, since all heteronomous principles ultimately are empirically conditioned, as they presuppose a contingent interest in some object (see e.g. Cohen 2014: 163). The charge of heteronomy, however, does not revoke the difference between empirical and rational principles, which Kant notably still used after introducing the heteronomy objection (see 5:41). The merely descriptive distinction between empiricist and rationalist views points at a fundamental difference in moral epistemology and ontology without anticipating any objection. The views that Kant calls rationalist remain so, even if the heteronomy objection applies to them, as long as they maintain that the moral

⁵ On the various 'histories of morality' in Germany prior to Kant see Hochstrasser 2004 (where, unfortunately, the remarks on Kant's own approach are insufficient; see Hochstrasser 2004: 201f).

⁶ The distinction is put in stronger terms in the Mrongovius II notes, where the key issue of modern views is: "On what, then, does morality rest?" (29:620). Cf. Kaehler 20.

standard builds on something that cannot be cognized within experience, such as metaphysical perfection or the perfect will.

My examination shall focus on Kant's comments on the views that he calls 'rationalist', which prove to be closely connected with the development of the idea of autonomy.⁷ Kant's treatment of the rationalist views is especially worth examining also because in the earlier published writings, such as the *Distinctness* essay of 1762/64, his remarks seem to privilege the sentimentalist views, apparently neglecting the 'rationalists'. Disregarding the role of Kant's discussion with the 'rationalists', however, amounts to ignoring a crucial aspect of the emergence of the idea of autonomy. The lecture notes, along with Kant's private notes, thus provide a highly valuable supplement to the published writings also in this regard.

Before examining Kant's criticisms of rationalist accounts a last clarification on his taxonomy is needed. Kant applies a further distinction between 'external' and 'internal' principles to both empiricist and rationalist views. This difference does not concern the subject and its relation to the action, as is generally assumed.⁸ What is at issue, is not whether the ground of moral demands is internal to moral subjects, but what is a proper account of the principle that determines the moral worth of actions. In fact, Kant here distinguishes between opposing conceptions of the moral quality of actions, which can be understood as internal or external *to the action itself* (see e.g. 27:109, 27:120). Kant draws on the distinction between *moralitas objectiva sive interna* and *subjectiva sive externa*. The former designates the inner worth of an action, understood as an intrinsic property of the action, while the latter is the worth of an action depending on a commanding subject's will.⁹ Like his contemporaries, Kant distinguishes between views that centre on "inner morality

⁷ Here I cannot examine to what extent Kant adequately represents the views he discusses. For a brief overview of most of them and references to further literature, see Bacin 2017a. On Kant's remarks on the 'empiricist' views see Georg Mohr's chapter.

⁸ See e.g. Schneewind 2009: 145; Kerstein 2002: 142. According to Wood 2015: 124, 'inner' here means 'arising from our own will', which cannot apply to the perfectionist view as Kant construes it. Wood accordingly understands the principle of education as inner, while Kant considers it as external (see Wood 2015: 124 and 29:621).

⁹ See Baumgarten 1760: § 36 and 82; Meier 1764: § 51, 103f: "Internal morality has its sufficient ground primarily in the very quality of the action itself and in the nature of that person who performs it. [...] External morality of a free action is that which has its sufficient ground primarily in the free will of a person". Cf. Wolff 1733: § 137, 392f and the definitions in Meissner 1737: 388.

[arising] from the nature of the action” and views that argue for “external [morality], from duty towards commands” (Refl. 6480, 19:21). A principle is external if it construes morality as having its ground “outside the nature of action, lying in the will of another” (Refl. 6754, 19:149). On the other hand, an internal principle construes morality as belonging to the action, without depending on the imposition of any will. Thus, in the rationalist camp, the perfectionist view chiefly represented by Wolff (cf. 29:626) is not described as internal because it would maintain that the source of morality is internal to the agent, as it might seem. This reading would not even match Wolff’s position, which does not make morality depend on any function of the subjects. Kant regards the principle of perfection as ‘internal’, instead, because it holds that the moral quality of actions is objective and intrinsically bound to essential features of the actions themselves. Analogously, the theological view championed by Crusius is construed as “external” (cf. 29:627) because it claims that the moral worth of actions depends on an external act of God’s will. Therefore Kant gives a rather traditional description, which arguably means to be uncontroversial.

Kant’s distinction between empiricist and rationalist views thus include in ‘rationalism’ principles that we would now call voluntaristic, as opposed to rationalist, as is still apparent in the *Groundwork* (4:443) and the second *Critique* (5:39).¹⁰ This is remarkable, because voluntarist and rationalist accounts of morality conflicted at least since Cudworth’s criticism of Hobbes (see e.g. Schneewind 2010). Kant nevertheless presents as rationalist* both the perfectionist *and* the theological views.¹¹ He justifies the inclusive description of the rationalist* camp referring to the idea of perfection as a unifying feature, maintaining that all the views falling in the rationalist* camp ultimately build on some idea of perfection. While this clearly holds for the perfectionist conception, the claim that the theological principle is centered on the idea of “a most perfect will” (29:627, cf. 5:41) entails that the divine will can be cognized only through reason, just like any notion of perfection. The theological view belongs to the rationalist* camp because it presupposes that God’s will is “known through

¹⁰ To prevent misunderstandings, I will hereafter use ‘rationalism*’ for what Kant calls ‘rationalism’ or ‘intellectualism’, and ‘rationalism’ for the current, narrower sense of the term.

¹¹ Kant sometimes also mentions a “principle of truth”, which he rather oddly attributes to Cumberland (cf. 27:121, 29:622). In a private note Kant even leaves a blank for the name of its advocate: cf. Refl. 6624, 19:116, and the editors’ note. Kant must have meant Wollaston’s view. See also Wood 2015: 125.

reason, and not positively” (27:510). Thereby Kant rules out from the beginning simpler versions of divine command theories, which are centered on the thought of the unconditioned sovereign power of God as entailing the capacity to reward and punish here and in the afterlife. According to Kant, such accounts should not be regarded as rationalist*, since God’s will would not be the object of a rational cognition, but rather of an empirical attitude of fear, submission, or gratitude resulting from a contingent positive act. Kant’s selective understanding of the theological paradigm is neither artificial nor motivated by mere taxonomical reasons. On the contrary, it focuses the examination of the theological account from the outset on the necessary connection of a divine command view with the intrinsic perfection of God’s will, thereby highlighting what emerges as its crucial feature.

While Kant’s charts of the rationalist* accounts of morality include a voluntarist view like the theological conception, they do not mention prominent rationalist views like Samuel Clarke’s or Richard Price’s. While he may or may not have known of Price’s views since the *Review of the Principal Questions of Morals* was never translated into German, French, or Latin, Kant must have known something of, or about, Clarke or Cudworth.¹² Since the aim of these remarks is not to provide an exhaustive map of the current views, however, the absence of some names does not entail that the criticisms are not meant to apply to further writers as well. Kant might arguably have understood those rationalist views as further instances of the principle of (metaphysical) perfection.¹³

2. The Arguments against the Principle of Perfection

Kant’s remarks against the perfectionist view of morality can be summed up in two main points, which are often repeated both in the lectures and in Kant’s notes.

A first criticism against the principle of perfection, which can be called the Contingency argument, points out that because human perfection is contingent, it cannot provide any normative self-standing criterion: “Solely from the fact that it is in accordance

¹² Clarke’s main works were available in German translation in Kant’s times: see Clarke 1756, which includes both cycles of Boyle lectures. On Cudworth see footnote 26. Interestingly, Kant’s moral view was early compared to Price’s and Cudworth’s: see e.g. Meiners 1801: 85-9 and Stäudlin 1822: 961.

¹³ For a different suggestion see Wood 2015: 125.

with our nature, it is not perfection merely, for I can have a better nature, e.g. angel" (27:6; cf. 27:62). This objection is especially relevant in the first phase of Kant's moral thought, in the mid-1760s, when he supposes that the perfectionist views could be supplemented with a thorough examination of human nature: "Hence the supreme law of morality is: act according to your moral nature" (27:6). The opposition between possible perfection and real perfection points out the risk that ethics be merely "chimerical", that is, unable to make demands that can be enacted. This is one important reason why Kant, in the earlier phase of his practical philosophy, suspects that reason cannot be the only faculty of moral cognition since it cognizes perfections higher than what human beings can possibly achieve (cf. 27:62).

Kant's second main criticism against the principle of perfection can be called the Emptiness objection (cf. 27:16, 29:627; Refl. 6754, 19:149). Prescriptions like "do good" (*fac bonum*) or "perfect yourself" (*perfice te*), which Wolffian perfectionism presents as the highest moral rules, are nothing but tautologies expressing the prescriptive meaning of moral terms ('good', 'perfect'), and do not determine any specific duties (see Bacin 2006: 86). Kant's observation concerning "*fac bonum*" applies to the command of perfection as well: "there ought to have been another rule, to tell us what moral goodness [or moral perfection] consists in" (27:264).¹⁴ Kant argues that the notion of perfection is empty, and thus cannot have any practical significance because it merely results by abstraction from actual cases of moral pleasure: "the general concept of perfection is not comprehensible through itself, and from it no practical judgments can be supplied; rather it is itself more a derived concept in which that which occurs in particular cases is given the general name 'perfect'" (Refl. 6624, 19:116; cf. Refl. 6634, 19:120). This objection is sometimes connected with the hint at a positive project, which follows from the thought that the notion of perfection can determine a proper moral principle once it is made more specific. Therefore, Kant proposes a distinction between moral and pragmatic perfection (cf. 27:130). If the former is construed as the perfection of the will, then "perfice te" can indeed be considered as a moral principle, as he once suggests, anticipating the *Doctrine of Virtue*.¹⁵

¹⁴ Guyer 2007: 306, argues that the many pages that Wolff devotes to the duties to oneself in his ethical works belie Kant's charge of tautology. Those detailed remarks, however, show not that commanding perfection can determine the contents of moral demands, but only that a number of moral demands can be presented as instances of the general demand of greater perfection.

¹⁵ Cf. 27:265. Guyer 2011 has pressed this point.

With regard to the philosophical debates of Kant's times, he notably never embraces more malicious criticisms like Johann Joachim Lange's Thomasian point that suggests Wolff's perfection is nothing but a metaphysical name for self-love and egoism. In fact, Kant already presupposes Baumgarten's distinction between perfection as a means and perfection as an end, meant as a response to that objection (see Schwaiger 2011: 135). When Kant comes closest to such an objection, is when he observes, not that the Wolffian principle is inherently non-moral, but that because of its vagueness it fails to determine a moral criterion, which amounts to the charge of emptiness.¹⁶

3. The Arguments against the Theological View

Kant's criticisms of the theological view are more extensive and differentiated than those directed against the principle of perfection. They narrow down to four main objections.

First, Kant argues that a theological account of morality inevitably leads to Arbitrariness and Contingency. According to this objection, if we understand morality as issued by God's will, as the theological accounts maintain, we thereby make morality contingent, which it cannot be. On this view, the moral laws become "*statuta*", i.e. positive laws (e.g. 29:627). On the contrary, Kant holds that it is an essential feature of moral norms to have the validity of 'natural laws', as opposed to positive laws. For him, the advocate of a theological view "regards natural laws as *statuta* of the divine will, but they are norms that lie from eternity in the nature of the thing" (27:173; cf. 27:546). The standard of morality cannot be freely determined by an author of the law, contrarily to what a theological voluntarism would entail (cf. 27:136, 29:633f). Kant repeats this argument often, mostly in

¹⁶ Cf. 27:264: "the statement *Fac bonum et omitte malum* can be no moral principle for obligation, for the good can be good in a variety of ways for any given purpose, since it is a principle of skill and prudence; only if it is good for moral actions would it then be a moral principle. It is therefore a *principium vagum*". The same point is still made in the later Vigilantius notes: cf. 27:517.

connection with the distinction between author of the law and author of the obligation, i.e. legislator.¹⁷

The Contingency argument, however, is not an original criticism. Leibniz, for instance, had made the same point against Pufendorf, arguing that “no one will maintain that justice and goodness originate in the divine will, without maintaining that truth originates in it as well: an unheard-of paradox”. On the contrary, Leibniz maintained that “justice follows certain rules of equality and of proportion no less founded in the immutable nature of things, and in the divine ideas, than are the principles of arithmetic and of geometry”.¹⁸ But Cudworth had also criticized voluntarism along the same lines (see Cudworth 1731: § II.1-2, 16f). The Contingency objection thus appears to be the most straightforward for Kant, since it follows from the assumption that moral norms must be necessary rather than contingent and are to be understood as “natural laws”, i.e. as non-positive normative principles. Yet, this is remarkably the only objection against the theological principle that Kant does not mention in the *Groundwork*.

Secondly, Kant claims that the theological principle needs to presuppose the validity of moral standards prior to God’s lawgiving (Goodness and Divine Goodness). However, if a notion of goodness is necessary to acknowledge God’s act as good, that notion cannot derive from God’s act itself. Already the Herder notes report Kant wondering: “Can we, even without presupposing God’s existence and His *arbitrium*, derive all obligations from within? *Responsio*: not merely in the affirmative, for this, rather, is *ex natura rei*, and we conclude from this to God’s choice” (27:9). If that necessary assumption is not made, it is impossible to recognize the perfection of God’s will (cf. 19:35: “If one wants to cognize God prior to morality, one ascribes to him no moral perfections”). This suggests that we can know God’s will only through the moral laws (27:135f), or “from the nature of good actions” (27:109). This objection corresponds to the first remark on the theological account in the *Groundwork*

¹⁷ See 19:131; 19:156; 27:247, 283; 27:1433; 27:544f; Kaehler, 79. On the distinction see Kain 2004, Reath 2013b: 37-40.

¹⁸ Leibniz 1706: 71; cf. Leibniz 1710: 34f, 398f. Remarkably, Kant never directly discusses the possible reply that addresses the assumption that God’s commands are arbitrary. Kant seems never to dispute that “the faculty of choice does not determine [*constituiret*] anything but what is in itself contingent” (27:147). Barbeyrac had challenged precisely this assumption to defend Pufendorf against Leibniz (cf. Barbeyrac 1735: 415).

("[...] merely because we cannot intuit its perfection but can derive it from our concepts alone, the foremost of which is that of morality": 4:443).¹⁹ Leibniz, however, had already argued this point against Pufendorf: "One must pay attention to this fact: that God is praised because he is just. There must be, then, a certain justice — or rather a supreme justice — in God, even though no one is superior to him" (Leibniz 1706: 71).

A third remark pinpoints the epistemic uncertainty that affects the theological view, as it entails the requirement of revelation. If the "theological moralists" do not hold that God's will can be cognized through purely rational moral notions, then they must appeal to some sort of revelation. But such kind of cognition of the divine will cannot be a necessary enabling condition of complying with moral demands (cf. e.g. 20:112, 29:627). Kant thereby reverses a point of Crusius', who had claimed that referring to God's will would account for the possibility of cognizing what morality demands without presupposing any further knowledge.²⁰

A fourth critical remark, closely connected to the Goodness objection, concerns the justification of God's moral authority (Unjustified Authority). If God's commands are not grounded on any antecedent foundations, then their authority only builds on his power to impose sanctions on his subjects, just like any human sovereign. If God's authority in issuing commands is not justified, then the only ground of such commands would be power, much like in the case of a tyranny. But this would make moral demands merely pragmatic (see 29:627). Since any relevant distinction between theological and political variants of voluntarism disappear, the same objection applies. This corresponds to Kant's second remark against theological views in the *Groundwork*.²¹ Again, Cudworth had made the same

¹⁹ This objection is still present in the later Vigilantius notes: cf. 27:498.

²⁰ See Crusius 1744: § 135 (trans. Schneewind): "coming to know God and his will through proofs, and judging what is told us by means of clear arguments, is for most people much too roundabout. Most scholars do not get far in doing so, and [so] how could the unlearned get things completely right? But because all men are subordinated to the divine law and will be judged by it hereafter, it is easy to infer that God would create a shorter path to knowledge of it and would make his will evident in such a way that it could come to everyone's knowledge."

²¹ Cf. GMS 4:443: "but because if we do not do this [i.e., if we do not derive our conception of God's will from our moral concepts] [...], the concept of His will still left to us – taken from the attributes of desire for honor and dominion, ... combined with horrendous representations of power and the zeal for vengeance – would have to be the foundation for a system of morals directly opposed to morality".

objection and Leibniz had used it against Pufendorf: “If the source of the law is the will of a superior and, inversely, a justifying cause of law is necessary in order to have a superior, a circle is created, than which none was ever more manifest”.²² Since the fear of punishment cannot be the foundation of morality, Kant argues that God’s act of will presupposes an obligation that is already in place (27:9).

A theological view prone to the Unjustified Authority objection, however, could hardly be regarded as rationalist* since the appeal to some perfection is here lacking. Accordingly, the cognition of moral demands cannot be considered rational. In fact, all of Kant’s remarks against the theological account ultimately press this very point, that is, the necessary connection between the divine will and perfection, which is already emphasized in regarding that view as rationalist*. If God’s will is to be regarded as the perfect will, it must never make an arbitrary choice (objection 1) because it operates according to the criterion provided by an inherent, independent goodness (objection 2). Therefore its workings are fully intelligible to every rational being without any revelation (objection 3) and its authority is justified by the intrinsic connection with the perfect goodness (objection 4). All objections imply that the theological view is erroneous insofar as it turns out to be an empirical account, namely as it builds on conditions that, unlike an idea of perfection, are only empirically cognisable.²³

4. A Rationalist’s Critique of Rationalism

Kant’s objections against rationalism*, however, lack remarks on the conception(s) of reason advocated by the writers considered, or implicit in their views. No specific conception of reason comes into play, nor are the features of the reason that Kant is talking about in labelling these principles ‘rational’ ever discussed.²⁴ This does not appear to be the

²² Leibniz 1706: 73 (cf. 73-5). Cf. Cudworth 1731: § II.3, 19.

²³ See the later Vigilantius notes: “if the ground of duty is posited in the divine will, it is either drawn from metaphysics, namely [from the] *theologia naturalis*, insofar as the divine will is known through reason, or else derived from *theologia revelata*, and in that case it is empirical” (27:538).

²⁴ See e.g. Guyer 2007 and 2011, where he argues that Kant’s new notion of reason is the key difference between him and Wolff, especially in moral philosophy.

main issue, for Kant, or at least not the right way to frame it. Kant's arguments are likely meant to hold independently from a specific notion of reason, as they merely presuppose a comparatively uncontroversial conception of the faculty of the mind enabling to cognize non-empirical truths. Furthermore, if Kant maintains that pure practical reason has a lawgiving role, this claim is part of his solution to the issue of the bindingness of moral demands. He does not directly object to rationalists* that their understanding of the power of reason is crucially limited, (a) because providing an accurate account of reason is not the aim of their moral views, and (b) because such argument would in fact presuppose Kant's own view.

Notably, Kant's critical remarks display no substantial changes through the various versions of the lectures. Most, if not all, objections are already present in the Herder notes and other texts of the 1760s. Furthermore, Kant mostly draws on objections previously put forward by Leibniz and other writers. If Kant's criticisms of the previous principles were meant as a step towards a more satisfying solution to the issue of the "supreme principle of morality", as I have suggested, then how do they contribute to the development of the idea of autonomy?

A way of understanding Kant's critical remarks could be to construe them as an 'only survivor' argument: "'All the other available views have fatal flaws', this kind of argument goes, 'and mine alone survives all the criticisms; so mine is the best'" (Schneewind 2009: 141). Something of the kind does surface in Kant's notes:

"I. The *principium* of morality is not sensual, neither *directe* or pathological, lying neither in the physical (doctrine of skill) nor the moral sense (the latter is impossible, because no sense intrudes on the intellectual); nor *indirecte* sensual or pragmatic (doctrine of prudence): considered according to your true happiness (Epicureanism). There reason serves only as the means to determine how the greatest sum of inclinations can be satisfied, and the means for this. The *principium* of morality is thus intellectual (pure), but not tautological (*perfice te, medium tene*). II. The *principium* is not external, outside the nature of action, lying in the will of another." (Refl. 6754, 19:149)

Passages like this, however, do not provide compelling evidence that Kant thought of his criticisms as a single argument by elimination for the principle of autonomy. Similar

passages are arguably *ex post* attempts to organize the presentation of the previous views in a critical summary like those given in the *Groundwork* and the second *Critique*. But this does entail that the function of Kant's criticisms is to highlight the only possible solution only by opposition. More importantly, however, interpreting those criticisms as an argument by elimination obscures two decisive, closely connected elements. First, Kant's stance to the previous principles is not as homogeneous as the 'only survivor' reading implies, since he does regard some principles as closer to the truth than others. Second, Kant tries to incorporate aspects of other positions in his account (see also Bartuschat 2004). I shall clarify these points in turn.

The two aspects are closely connected because a first intermediate result of Kant's scrutiny is that empiricist principles are inadequate for explaining the bindingness of moral laws, while the rationalist* accounts are more promising (see also Wood 2015: 124). Therefore, a new account of the principle of morality should attempt a combination of the two rationalist* views. From very early on, Kant sees them as a more promising attempt at an adequate account of morality. His understanding of the advantage of rationalist* conceptions is especially clear when he presents the empirical principles as based on "contingent grounds", while the goal is to explain morality "from the nature of free actions" (27:106). This is precisely what rationalist* views try to do, especially the internal view (cf. 29:622). More exactly, Kant holds that only a rationalist* view can account for what he considers as an essential feature of morality, that is, its necessity:

"The empiricism of morality proves only *that no one approves of a lie; rationalism, however, that no one can approve it*, and indeed for itself alone; hence only in the latter case are moral imperatives apodictic. The cause is: because a freedom without rule taken as universal authorization contradicts itself. The system is thus a rational system of freedom in universal consensus with itself" (Refl. 7217, 19:288).

For Kant, here lies the general difference between empiricism and rationalism. As he puts it in the second *Critique*, "empiricism is based on a necessity felt, but rationalism on a necessity seen [*eingesehen*]" (5:13). Only rationalist* views maintain the necessity and objectivity of morality (see especially Kaehler 26, 27:254, 27:120).²⁵ This divide is so clear, for Kant, that the view that he is searching to articulate is often just listed as one of the

²⁵ On this assumption in pre-Kantian rationalists: see Gill 1998.

rationalist* accounts.²⁶ Kant holds that, since the available rationalist* views are not successful, a novel *rationalist* account of morality is needed.

Thus, Kant's critical remarks on the rationalist* views do not constitute part of a continuous argument, but instead examine unsatisfactory answers to the question at issue that are, however, instructive in important respects. An 'only survivor' argument would entail that, one after another, all examined options are fully discarded. But understanding Kant's search in those terms obscures the fact that his new solution, centered on the idea of autonomy, is not only growing out of the rationalist* camp, but is also tied to *both* rationalist* views. To properly understand Kant's stance, it is necessary to look into his view of the relationship between the two main variants, the perfectionist and the theological.

5. Rationalism and the Law-Giving Will

Kant does not discuss the perfectionist view very extensively, and gives greater prominence to the arguments against the theological account. This happens for two reasons: first, the importance of theological accounts in the moral philosophy of his times, and second, the highly significant 'grain of truth' contained in the theological views.

Kant regards the theological principle as the most fashionable position of his time. His special attention to it seems in fact to be at least partially motivated by a worry regarding its success. In the *Mronovius II* notes, he remarks that, while "the principle of morality from external empirical grounds has but few adherents" and "the metaphysical principles [namely, of perfection and truth] are nowadays largely abandoned", "a majority has fallen back on the theological principle, because the metaphysical one lacks force" (29:622). The

²⁶ See Kaehler 27 (cf. 27:255): "what this inner intellectual principle consists in, it will be our aim in morals to determine, though it can only be extracted gradually, over time." Note that the label "systema intellectuale" that Kant often uses for the rationalist* views might suggest a connection with Cudworth. Cudworth's main works were available in Samuel Mosheim's Latin translation under the general title *Systema intellectuale hujus Universi* (Jena 1733; reprint Leiden 1773), which includes, besides the *True Intellectual System*, the *Treatise Concerning Eternal and Immutable Morality*, in which Cudworth states many of the arguments Kant deploys against the theological view. Kant must have been acquainted with Cudworth's works, since he owned Mosheim's edition (see Warda 1922: 47; also Tonelli 1967: 117 n. 29; Kuehn 2001: 470 n. 19).

mention of Crusius along with “other theological moralists” in the second *Critique* (5:40) also suggests that Kant’s worry does not just concern a few writers, but is a more general criticism. Besides Crusius himself and his followers, a theological view along similar lines had been defended also in Königsberg in the very years when Kant began to work on these matters. Daniel Weymann, Kant’s nemesis in the early years of his career, as it were, published in 1762 (roughly at the time when Kant was working out his first thoughts on the foundations of morals) a brief dissertation *De vero stabiliendo juris naturae et gentium principio* (*On Establishing the True Principle of Law of Nature and Nations*), which, under a Pufendorffian sounding title, argued for a broadly Crusian view of morality.²⁷ A theological conception was brought once more to the forefront by Barbeyrac, whose French translations of Pufendorf’s main works provided an influential support to the divine command view, along with responses to earlier criticisms (see Schneewind 1996 and Korkman 2003). Finally, Kant also included in the theological camp the very authors he lectured on. Not only did Achenwall hold such a view, as Kant remarks (cf. 27:1334), but Baumgarten also (and his follower Meier) had a similar position. While Wolff put greater emphasis on the naturalistic elements of his view, Baumgarten’s account had a distinctive theological turn, which entails some quasi-voluntarist features:

“the author of the obligation announced by the law is said to declare that law, and he who has the right to declare the laws is said, in an extended sense, to be a legislator [...]. Now God is the author of the whole nature, and of all realities that come to pass, and moreover all natural obligations are something real and positive, and have sufficient reason in the same. *Therefore God is the author of obligation, and thence also of the natural law*” (Baumgarten 1760: § 100).

Kant stresses this aspect of Baumgarten’s position, expressing his disagreement in the clearest terms: “Our author says: moral laws can be viewed as positive divine laws, but that is wrong” (29:635). He even points out that, in this respect, Baumgarten’s and Crusius’

²⁷ See Schwaiger 2011: 146. On Kant’s relationship to Weymann, see Kuehn 2001: 122-4.

positions are indeed not far apart.²⁸ In fact, in his lectures Kant introduces the important distinction between the author of the law and the legislator as author of the obligation of the law precisely to rectify this point of Baumgarten's (see Bacin 2013: 57).

The theological principle thus appeared to be the most prominent competitor in the discussion on the foundations of morals. In spite of fundamental mistakes, however, the theological principle provides a momentous clue that demands careful appreciation. On the one hand, Kant declares that the principle of perfection, with all its shortcomings, is undoubtedly superior to the theological principle (cf. 4:443).²⁹ Similarly to Leibniz, who regarded it as a "repugnant" and morally dangerous conception (Leibniz 1706: 72), Kant once even calls the theological view "altogether wrong and false and deleterious" (27:135). On the other hand, though, for him the principle of perfection does not provide much material for further elaboration. Perfectionism merely states, in regrettably vague terms, one central requirement. Kant thus observes that it is "not at odds with morality, but it cannot contribute anything to it either" (29:626). For this reason, it does not require a very extensive examination, since the Emptiness (or Vagueness) objection settles the question.

As perfectionist rationalism rightly acknowledges, the status of moral laws cannot allow for an author that *ex nihilo* determines their normative content. This, however, does not settle the question of obligation, since it lacks an explanation for the bindingness of moral laws. For this, they need the lawgiving act of the author of their obligation. In the perfectionist view, Kant implies, there is no conceptual space to argue for that. This is possible, however, if the explanation focuses on the lawgiving act of a rational will. This claim is the distinctive grain of truth contained in theological accounts, which made them worthy of special attention for Kant. While he adopts the arguments of previous critics of theological voluntarism like Leibniz, he never concludes from the rejection of the theological

²⁸ See 27:510: "Although the obligation is established by reason, it is nevertheless assumed that in the performance of our duty we have to regard ourselves as passive beings, and that another person must be present, who necessitates us to duty. Crusius found this necessitating person in God, and Baumgarten likewise in the divine will, albeit known through reason, and not positively, and on this principle a particular moral system has been erected." On Kant's contrast to Baumgarten on this matter, see Bacin 2015: 24-9.

²⁹ Stern 2012: 57 (and *passim*) greatly stresses this point, interpreting theological voluntarism as the main target of Kant's objections. On my reading, this is only one side of the coin.

view that morality does not essentially need prescriptions and that commands are only a sort of backup, since the rational agent does not will according to rules.³⁰ Other rationalists had explained moral obligation by arguing that the intrinsic “fitness” of some actions directly imposes an obligation, without requiring any act of will.³¹ On the contrary, Kant implies that such attempts collide with the conceptual impossibility of a genuine obligation for finite beings without reference to any will.

In Kant’s eyes, the theological view oscillates between the errors that make it equivalent to an empirical account and the grain of truth that makes it a more complex case of a perfection-centered view. In the latter case, the grain of truth prevails, but the theological element ultimately dissolves, since the decisive factor lies in the connection between the rational will as such and the idea of a necessary moral law, and such a connection can also be argued for with regard to finite rational beings. While Barbeyrac insisted that God’s lawgiving cannot be regarded as arbitrary, as it is bound to the criteria of rationality (see fn. 17 above), Kant presses the same point with regard to every reason. As a telling passage in the Powalski notes puts it, advocates of a theological account derive moral perfection “either from experience or a priori”. Only in the latter case does the theological view properly belong to rationalism*. In order for the appeal to God’s will to make sense, the “theological moralists” must refer to an independently valid standard: “They must derive morality *and* the divine will from the inner quality of actions”. However, the moral worth of actions can be cognized by reason as such (“understanding can derive the good and bad from morality”). Therefore, it is ultimately not necessary to refer to the *divine* will: “Since the understanding is the *principium* of morality, the theological principle drops out” (27:121). The best version of the theological account would thus no longer require *God’s* will as its foundation, since it revolves around the capacity of a rational will that, as such, draws on a necessary moral standard generating binding demands through universal volitions.

³⁰ See e.g. Leibniz 1706: 72: “Whoever does good out of love for God or for his neighbor, takes pleasure precisely in the action itself (such being the nature of love) and does not need any other incitement, or the command of a superior; for that man the saying that the law is not made for the just is valid”.

³¹ See e.g. Clarke 1706: 626f. Wolff’s notion of natural obligation is much along similar lines. See Wolff 1720: § 9 (trans. Schneewind).

Reworking this thought with respect to every rational will is a decisive step in Kant's development of his conception of the autonomy of the will. One remarkable passage in the *Mrongovius II* notes suggests that, in spite of the weaknesses, the theological principle points out the necessity of a reference to a will in explaining moral obligation. After presenting the objections that I have examined above, Kant there remarks:

“It seems as if duty had as its foundation the will of a legislator, not something one does according to one's own will but according to the will of another. But the will of another is not the will of another being; but rather our own will, *in so far we make it universal* and consider it as a universal rule. Such a will proceeds as a *universal*, not as a private will” (29:627, my emphasis).

Kant's dissatisfaction with the inability of perfectionist rationalism to account for moral obligation remarkably matches the critical comment on Wolff's universal practical philosophy in the *Groundwork* (4:390; cf. 29:598). Universal practical philosophy can be regarded, from Kant's standpoint, as the most detailed attempt to include in practical philosophy a investigation into the will.³² Such an attempt cannot but miss the point, though, if it does not explore the intrinsic connection between will and obligation. In Kant's view, universal practical philosophy only considered a generic will as the *addressee* of obligations, that is, according to Wolff's view of obligation, a capacity to act according to motives.³³ Therefore he missed the crucial link between obligation and will, namely its *active* side, because he did not consider under which conditions, on which principles, and thanks to which capacities, a will can *issue* a valid moral obligation. The theological view provides an insight that allows to fill this lacuna. The underlying thought in Kant's incorporation of the substance of the theological view, thus, is not to substitute the divine sovereign will with a God-like 'transcendental subject' who declares the unconstrained legislative power of his will, as has often been objected. On Kant's view, the theological view is significant, instead, because it gives prominence to the role of a will as author of moral obligation, pointing at a

³² Note that Kant's criticism also applies to empiricist versions of universal practical philosophy like Johann Georg Heinrich Feder's. See Bacin 2017b.

³³ Wolff understands obligation as the connection of an action with a motive. See Wolff 1720: § 8, (trans. Schneewind): “To obligate someone to do or omit something is only to connect a motive of willing or not willing to it”.

rational will that operates under the constraint of an independent standard of morality in universal volitions.³⁴ While Kant follows the rationalist tradition in embracing the Pauline thought of “being a law unto oneself”, he construes it differently than Clarke, Price and Wolff, and suggests that that thought conveys the decisive property of the lawgiving *will* (see GMS 4:440, 29:628).³⁵

Kant’s remarks on the two main rationalist* principles, thus, ultimately suggest a combination of them as the adequate solution of the issue of moral obligation. Different, yet apparently complementary advantages of the two rationalist* principles already surfaced when Kant was strongly interested in sentimentalist views. In the *Distinctness* essay of 1762/64 he observed that the general rule “do the most perfect thing”, which corresponds to the perfectionist standard of morality, provides the formal principle of morality. However, no specific obligations derive from it. For this, “indemonstrable material principles” are needed (2:299). Kant’s only example of such a principle is none other than the theological principle: “do what is in accordance with the will of God” (2:300). The early essay did not articulate how the suggestion should be worked out. In fact, a longer process was still required to formulate Kant’s own solution. Nevertheless, the early remarks show that Kant considers the two rationalist* views as compatible. He implies that they must be integrated, in order to properly explain both the universal form of moral obligation and the bindingness of specific demands. All this suggests that, at least since the early 1760s, one of the main goals of Kant’s work in moral philosophy was to find the right way to combine the insights of the two rationalist* principles. Looking for the appropriate combination represents a seminal thread in his development of the idea of autonomy.

6. Conclusion

In the *Groundwork* and the second *Critique* the previous views of morality are only discussed with regard to the general charge of heteronomy, which, derived by contrast from

³⁴ Kant’s conception could arguably draw as much on Malebranche’s theory of God’s general volitions as on Rousseau’s *volonté generale*, as is mostly assumed. See Riley 1986.

³⁵ See Andy Reath’s chapter in this volume. On the difference of Kant’s use of the Pauline thought from Butler’s, Price’s, and Wolff’s see Bacin 2013.

the new idea of autonomy (see 29:629), points out that all previous accounts miss that crucial thought. Before 1785, when the heteronomy objection is absent, Kant sketches a more differentiated assessment, providing significant insights in the development of his own conception.

The idea of autonomy introduced in 1785, in the *Groundwork* and in the contemporary lectures on natural law and moral philosophy, represents the outcome of a process in which the assessment of previous views has an essential part. Rationalist* conception, in particular, provided Kant with decisive conceptual elements to work out. When Kant puts forward autonomy as the supreme principle of morality (GMS 4:440; cf. KpV 5:33, 27:1326, 29:626), he presents a feature of the rational will – that is, its being “a law unto itself” – as the foundation of moral obligation. This solution, unthinkable for traditional rationalists, grows out of the suggestion that the advantages of the two rationalist* conceptions are complementary and must be integrated.

Kant’s project emerges from the failures of previous rationalist attempts and sets as its goal to work out a rationalist view that, by integrating insights from the theological conception, can account for the bindingness of moral demands. The new idea of autonomy emerges, in this respect, from the spirit of rationalism. Kant’s criticisms of rationalist* views thus do not merely assess the state of the art, but prepare the stage for a new position. Examining these accounts, he makes it clear that only a rationalist view can provide a satisfactory account of the necessity and categoricity of moral demands, since only reason can cognize necessary a priori concepts. At the same time, he highlights that the binding force of morality requires the act of a will. A combination of perfectionism and theological views thus provides the basic outline of the idea of autonomy of the will. Kant’s criticisms of the two rationalist* views thus leads to autonomy being construed as a two-tiered notion, which emphasizes both rationalist and voluntarist elements.³⁶ Kant’s assessment of

³⁶ On Kant’s idea of autonomy as a two-tiered notion combining rationalist and voluntarist elements see Bacin 2013 and Bacin 2017c.

rationalist* views recommends an account of autonomy that accommodates both the lawgiving function of a rational will and the legislation of a non-positive, necessary law.³⁷

References

Bacin, Stefano 2006. *Il senso dell'etica. Kant e la costruzione di una teoria morale*. Bologna: il Mulino

—: 2013. “Legge e obbligatorietà: La struttura dell’idea di autolegislatzione morale in Kant,” *Studi Kantiani* 26, 55-70

—: 2015. “Kant’s Lectures on Ethics and Baumgarten’s Moral Philosophy,” in Denis and Sensen 2015, 15-33

—: 2017a. “Rationalism and Perfectionism,” in *The Cambridge History of Moral Philosophy*, ed. Sacha Golob and Jens Timmermann. Cambridge University Press

—: 2017b. “Kant and Feder on the Will, Happiness, and the Aim of Moral Philosophy,” in Dyck and Wunderlich 2017

—: 2017c: “Moral Realism by Other Means: The Hybrid Nature of Kant’s Practical Rationalism,” in *Moral Realism or Anti-Realism in Kant*, ed. Elke Schmidt and Robinson dos Santos. Berlin–Boston: de Gruyter

Bartuschat, Wolfgang 2004. “Kant über Grundsatz und Grundsätze in der Moral,” *Jahrbuch für Recht und Ethik* 12, 283-98

Baumgarten, Alexander Gottlieb 1740. *Ethica Philosophica*. Halle: Hemmerde (second edition 1751, third edition 1763) [Second and third editions reprinted in AA 27:737-869, 27:873-1028]

—: 1760. *Initia philosophiae practicae primae*. Halle: Hemmerde [Reprint in AA 19:7-91]

³⁷ I should like to thank the other authors of this volume, most especially Andy Reath, for helpful comments on earlier versions of my chapter, and Jens Timmermann for making his revised text and translation of *Moral Mrongovius II* available to me.

Clarke Samuel 1706. *A Discourse concerning the Unchangeable Obligations of Natural Religion, and the Truth and Certainty of the Christian Revelation*, in *The Works of Samuel Clarke*, vol 2. London 1738

Cohen, G.A. 2014. *Lectures on the History of Moral and Political Philosophy*. Princeton University Press

Crusius, Christian August 1744. *Anweisung vernünftig zu leben, Darinnen nach Erklärung der Natur des menschlichen Willens die natürlichen Pflichten und allgemeinen Klugheitslehren im richtigen Zusammenhange vorgetragen werden*. Leipzig: Gleditsch, third edition 1767 [Reprint: ed. Giorgio Tonelli. Hildesheim: Olms 1969]

— 1745. *Entwurf der nothwendigen Vernunft-Wahrheiten, wiefern sie den zufälligen entgegen gesetzt werden*. Leipzig: Gleditsch [Reprint: ed. Giorgio Tonelli. Hildesheim: Olms 1969]

Cudworth, Ralph 1731. *A Treatise Concerning Eternal and Immutable Morality; with, A Treatise of Freewill*, ed. Sarah Hutton. Cambridge: Cambridge University Press, 1996

Gill, Michael 1998. "On the Alleged Incompatibility between Sentimentalism and Moral Confidence," *History of Philosophy Quarterly* 15, 411-40

Guyer, Paul 2000. *Kant on Freedom, Law, and Happiness*. Cambridge University Press

—: 2007. "Perfection, Autonomy, and Heautonomy: The Path of Reason from Wolff to Kant," in *Wolff und die europäische Aufklärung: Akten des 1. Internationalen Wolff-Kongresses*, ed. Jürgen Stolzenberg. Hildesheim: Olms, 299-322

—: 2011. "Kantian Perfectionism," in *Perfecting Virtue. New Essays on Kantian Ethics and Virtue Ethics*, ed. Lawrence Jost and Julian Wuerth. Cambridge University Press, 194-214

Hochstrasser, T. J. 2004. *Natural Law Theories in the Early Enlightenment*. Cambridge University Press

Kain, Patrick 2004. "Self-Legislation in Kant's Moral Philosophy," *Archiv für Geschichte der Philosophie* 86, 257–306

Kerstein, Samuel J. 2002. *Kant's Search for the Supreme Principle of Morality*. Cambridge University Press

Korkman, Petter 2003. "Voluntarism and Moral Obligation: Barbeyrac's Defence of Pufendorf Revisited," in *Early Modern Natural Law Theories*, ed. T. J. Hochstrasser and Peter Schröder. Dordrecht: Kluwer, 195–225

Leibniz, Gottfried Wilhelm 1706. "Monita quaedam ad Samuelis Pufendorffii Principia," in *Opera Omnia*, ed. Louis Dutens. Genève: Fratres de Tournes 1768, vol. 4.III, 275-28. Translation in Leibniz. *Political Writings*, ed. Patrick Riley. Cambridge University Press 1972, 64-75

—: 1710. *Essais de Théodicée sur la bonté de Dieu, la liberté de l'homme et l'origine du mal*, in *Die Philosophischen Schriften von Gottfried Wilhelm Leibniz*, ed. Carl Immanuel Gerhardt. Berlin: Weidman 1875–90, vol. VI [Reprint: Hildesheim: Olms 1978]

Meissner, Heinrich Adam 1737. *Philosophisches Lexicon aus Christian Wolffs sämtlichen deutschen Schriften*. Bayreuth: Vierling

Reath, Andrews 2013b. "Kant's Conception of Autonomy of the Will," in Sensen 2013a, 32-52

Reich, Klaus 2001. *Gesammelte Schriften*, ed. Manfred Baum. Hamburg: Meiner

Riley, Patrick 1986. *The General Will before Rousseau: The Transformation of the Divine into the Civic*. Princeton, N.J.: Princeton University Press

Schmucker, Josef 1961. *Die Ursprünge der Ethik Kants in seinen vorkritischen Schriften und Reflektionen*. Meisenheim: Hain

Schneewind, J.B. 1996. "Barbeyrac and Leibniz on Pufendorf," in *Samuel Pufendorf und die Europäische Frühaufklärung. Werk und Einfluss eines deutschen Bürgers der Gelehrtenrepublik nach 300 Jahren (1694–1994)*, ed. Fiammetta Palladini and Gerald Hartung. Berlin: Akademie Verlag, 181–9

—: 1998. *The Invention of Autonomy: A History of Modern Moral Philosophy*. Cambridge University Press

—: 2009. "Kant Against the 'Spurious Principles of Morality,'" in Timmermann 2009a, 140-58

—: 2010. "Voluntarism and the Foundations of Ethics,'" in *Essays on the History of Moral Philosophy*, Oxford: Oxford University Press, 202-21

Stern, Robert 2012. *Understanding Moral Obligation: Kant, Hegel, Kierkegaard*. Cambridge University Press

Timmermann, Jens 2015: "Mrongovius II: A Supplement to the *Groundwork of the Metaphysics of Morals*," in Denis and Sensen 2015, 68–83

Tonelli, Giorgio 1967. *Kant und die antiken Skeptiker, in Studien zu Kants philosophischer Entwicklung*. Hildesheim: Olms, 93-123

Warda, Arthur 1922. *Immanuel Kants Bücher. Mit einer getreuen Nachbildung des bisher einzigen bekannten Abzuges des Versteigerungskataloges der Bibliothek Kants*. Berlin: Breslauer

Wolff, Christian 1720. *Vernünfftige Gedancken von der Menschen Thun und Lassen (Deutsche Ethik)*. 4th edition. Frankfurt, Leipzig: Renger 1733 [Reprint: ed. H. W. Arndt, Hildesheim: Olms 1996.]

Wood, Allen W. 2015. "Kant's History of Ethics," in Denis and Sensen 2015, 120-37