Sabina Vaccarino Bremner
Unpublished draft as of September 25, 2019

Kant on Autonomy as Self-Making

1.1 The 'application presupposition' in Kantian conceptions of autonomy

Most Kant scholars have taken the Kantian principle of autonomy given in the *Groundwork* (GMS)[1] (1785) to be indispensable to Kant's mature moral philosophy (principally the *Metaphysics of Morals* (MS), 1797):

> The principle of autonomy is thus: not to choose in any other way than that the maxims of one's choice are also comprised as universal law in the same willing. (GMS 4:440)

On the standard view, the principle of autonomy is to be applied to MS (what I will call the 'application presupposition'), despite the fact that 12 years passed between the publication of the two texts, during which Kant's views underwent crucial changes, including the publication of the *Critique of Judgment* (KU).

Recent scholarship has challenged the application presupposition by suggesting that the principle of autonomy "virtually disappears" from Kant's later moral philosophy, a dramatic shift occasioning "no debate" in the literature (Kleingeld 2017: 61). While I agree that Kant's account of autonomy undergoes a change from GMS to MS (and that this development has so far not been appreciated in the literature), in this chapter I will argue that Kant *does* continue to invoke autonomy in his mature views—but in the *Opus Postumum* (OP) rather than MS. Autonomy, as it is now presented in OP, is no longer identical to the standpoint in GMS of the ideal moral legislator freely conforming to the categorical imperative; instead, autonomy is now conceived as a unified principle of 'self-making' that straddles both the theoretical and practical spheres. Although it ultimately aims at morality, autonomy as self-making therefore does not take an exclusively moral form. Moreover, it explicitly pertains to the empirically situated human subject, characterizing her active process of self-constitution as a rational being. Thus, while the principle of autonomy does 'virtually disappear' from MS as we had known it, its *re*appearance in OP has been neglected.

While Kleingeld argues that, despite the change in status of autonomy, Kant's

---

[1] References to Kant are to the volume and page of *Kants gesammelte Schriften* (except citations to KrV, which are to the A and B pagination) and accord with the following abbreviation scheme: ApH: *Anthropology from a Pragmatic Point of View*; B: *Correspondence;* DV: *Doctrine of Virtue* (part II of *Metaphysics of Morals*); GMS: *Groundwork of the Metaphysics of Morals*; KpV: *Critique of Practical Judgment*; KrV: *Critique of Pure Reason*; KU: *Critique of the Power of Judgment* (KaU: Part I, *Critique of Aesthetic Judgment*; KtU: Part II, *Critique of Teleological Judgment*); MAN: *Metaphysical Foundations of Natural Science*; OP: *Opus Postumum*; MS: *Metaphysics of Morals*; RGV: *Religion Within the Bounds of Reason Alone*; VA: *Lectures on Anthropology*; VE: *Lectures on Ethics.*

account of morality stays largely the same, I attribute the shift to a larger alteration in the conception of his philosophical system. In fact, Kant explicitly notes the influence of OP in the text of the *Doctrine of Virtue* (DV), in a passage (to which I return below) in which he acknowledges the need for a parallel 'transition' in both texts between pure principles (whether metaphysical or practical) and empirical cases.[2] In OP, this transition is secured by Kant's new doctrine of 'self-making', which he terms a form of autonomy; in MS, this transition is secured by practical judgment or casuistry, which applies pure moral principles to contingent circumstances. I argue that the conception of autonomy we get from OP—despite some important differences—is still relevant to the account of practical judgment Kant stresses throughout DV. As OP demonstrates, the new need for a schematism between pure moral maxim and contingencies in application leads Kant not just to *supplement* his initial conception of autonomy with accounts of moral training or moral character, as Kantians have often interpreted Kant's mature position,[3] but instead to expand the notion of autonomy to incorporate what Kant refers to as 'heautonomy' in KU, or the self-referential normativity of judgment: the *reflective* dimension of practical deliberation.

Given that existing accounts of OP focus almost exclusively on Kant's post-critical theoretical philosophy,[4] the notion that this work could also help to elucidate Kant's post-critical practical philosophy has, to my knowledge, remained largely unaddressed.[5] Yet devoting attention to the latter is of direct consequence to current debates on autonomy in political and moral philosophy, where the Kantian conception of autonomy has often been set in opposition to contemporary accounts. Joseph Raz, for example, takes pains to distance his conception of personal autonomy as "self-creation" from the "only very indirectly related" Kantian conception, since for Kant "authorship reduced itself to a vanishing point as it allowed only one set of principles which people can rationally legislate and they are the same for all" (1986: 370). David Johnston concurs that the appropriate conception of autonomy for the purposes of political theory is based on an ideal of "self-authorship" or "self-fashioning", which he contrasts with Kantian moral autonomy (1994: 75).[6]

---

[2] Thorndike (2018) is, to my knowledge, the only commentator who has acknowledged Kant's own stated affinity between the two works.

[3] See, e.g., Herman (2007): 133, Hill (1991): 43.

[4] See, e.g., Friedman (1992), Förster (1989, 1993b, 2000, 2012).

[5] For a recent exception, see Thorndike (2018), who—despite advancing an otherwise insightful reading of the connections between OP and Kant's practical philosophy—refrains from noting the new sense given to *autonomy* in OP.

[6] For a contrasting view that rejects such personal autonomy theories from a Kantian perspective, see Flikshuh (2013).

Christine Korsgaard has attempted to bridge the gap between these two conceptions by advancing a Kantian account of self-constitution. Her account makes space for individual variation and contingency by attributing it to differences in the *practical identities* individuals adhere to (such as 'mother' or 'student'). A conception of one's practical identity is "a description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking" (Korsgaard 1996: 101; see also 2009). However, Korsgaard does not claim to derive this account from Kant; she treats it as needed to supplement the Kantian account of moral autonomy with a source of the more contingent normativity that structures everyday life.

The alterations undergone by Kant's conception of autonomy in his final decade may show a way of reframing the presumed opposition between 'Kantian autonomy' and the contingent normativity of self-constitution or self-fashioning. First, in OP Kant suggests that autonomy can serve as a more general normative criterion than the employment of the categorical imperative; it also characterizes the subject's 'self-making' of her own judgment in thinking experience in accordance with norm-guided rational and moral ideas. Second, this development in Kant's account of autonomy helps to bridge the gap between universalizable maxims and particularities of practical reasoning, including individual variance in how one opts to structure one's own selfhood and course of life, which cannot be fully captured by a universalization procedure. Moreover, it does so not through a commitment to the notion of *identity*, which fails to allow for the questioning of the identities we inherit and that can fail to fit, but by appeal to Kant's notion of *reflectivity,* which allows both for the creation of new principles and the critique of preexisting ones.

This chapter proceeds as follows: first, I present my reading of Kant's account of autonomy as self-making in OP, situating this account within his general aims in this work. Then, I present grounds for why autonomy as self-making can be seen to constitute a significant shift from Kant's better-known conception of autonomy as self-legislating the moral law. I draw attention to Kant's claim in the *Doctrine of Virtue* that *both* a metaphysics of morals *and* a transition from metaphysics to natural science (the stated aim of OP) require a schematism. If this is the case, then one proposal I advance here for the role played by Kant's alternate conception of autonomy in OP is that it incorporates Kant's conception of the autonomy (or self-reflexive normativity) *of judgment*—a notion christened 'heautonomy' in the *Critique of Judgment.*

1.2   Autonomy in the *Opus Postumum*

1.2.1   *Self-making as a transition from metaphysics to physics*

To contextualize my discussion, I will start by giving a brief summary of Kant's general aims in OP.[7] As Förster (1989b) and Friedman (1992) have argued, Kant's central problem in this text is that of securing a transition between the metaphysical foundations of natural science (already defended in Kant's 1786 text of this title) and physics. Kant conceived of this task as fulfilling a central function for his critical system as a whole: in letters from 1797 (the year of publication of MS) and 1798, Kant described the prospect of such a transition as the "final goal" that would pay "the unpaid bill of my uncompleted philosophy", filling "a gap that now stands open" (12:222, 12:257, 12:258).

In order to secure this transition, Kant argues that we can examine our own experience as human subjects: simultaneously organic, moving bodies subject to the same laws of physics as other natural objects, and thinking beings who *create* the form of experience in which such laws are cognizable. If "experience must be made" (22:322), and "we have insight into nothing except what we can make ourselves" (22:353), in order to make sense of the transition between metaphysics to physics, "first… we must make ourselves": that is, first we must generate our own self-consciousness, or experience of ourselves as both subjects and objects of thought. Through the "act of composition [or synthesis] (*synthetice*)" of the manifold of self-consciousness, "the subject makes itself, according to a principle, into an object as it appears to itself" (22:358). In doing so, it evidences the fact that physical laws are ones to which we are subject in order for cognition to be possible: "The subject in appearance, which collects the inner moving forces for possible experience (for the completeness of possible perceptions) in conformity with a formal law… affects itself according to a principle, hence appears to itself as compositive (by inner moving forces)" (22:359). On the other hand, however, these laws are ones that we ourselves create in the act of cognizing, since "one is oneself the originator of one's own power of thought" (22:79).

It has been established that this position constitutes a departure from Kant's critical theoretical philosophy.[8] As these commentators note, Kant's post-critical views go in a much more substantially idealist direction—one comparable to that of his German Idealist interlocutors, such as J.S. Beck, Fichte, and Reinhold.[9] Like these figures, Kant, too, develops a *Selbstsetzungslehre* (theory of self-positing or self-composition) in OP. Indeed, Förster argues that Kant's *Selbstsetzungslehre* is not influenced by such figures

---

[7] Devoting more attention to this point is outside the scope of my paper; for more detailed accounts of Kant's aims in OP, see Förster (1987, 1989, 1990, 1993a) and Friedman (1992).
[8] See, e.g., Förster (1990), Wallner (1984), di Giovanni (1985).
[9] See also Kant's 1797 and 1798 letters to Tieftrunk (12:207 and 12:241), Garve (12:255), and Kiesewetter (12:258), in which Kant attacks these figures' popular new extensions of his critical project while jealously requesting news of how these new extensions are being received.

(particularly Fichte), but in fact precedes their own *Selbstsetzungslehren* (1989: 218). The OP, then, presents a 'common root' uniting the two heterogeneous faculties, one absent during the critical period: the activity of self-positing.[10]

As the OP progresses, the principle of self-positing (*selbstsetzen*) or self-affection (*sich afficiren*) begins to take the form of self-*making* (*sich machen*). It is the notion of self-making, developed on the basis of the possibility of self-affection, on which Kant lays emphasis in order to establish continuity between the theoretical and moral domains. In the realm of theoretical cognition, Kant argues that the fact "that there is something outside me is my own product"—that is, that experience is something that must be produced by me, the cognitive subject (22:82). By extension, however, *I* am my own product, since I must also produce my experience *of myself*: "I am an object of myself and of my representations… I make myself" (22:82). As a result, both internal and external experience, experience both of myself and the world outside myself (and thus 'everything'), is a product of my own making: "We make everything ourselves" (22:82).

The notion of self-making is then extended to the moral domain: "The subject determines itself (1) by technical-practical reason, (2) by moral-practical reason, and is itself an object of both" (22:53, 213). Indeed, Kant posits a causal connection between one's recognition of one's own powers of self-affection and one's recognition of one's own status as a *moral* being:

> According to transcendental idealism, the subject constitutes itself *a priori* into an object—not as given in appearance… but as a being who is founder and originator of his own self, by the quality of personality: the '*I am.*' As a man, I am a sense-object in space and time and, at the same time, an object of the understanding to myself. *I am a person;* consequently, a moral *being who has rights*. (21:14, 221).

The notion of being one's own originator or maker is what constitutes the basis for how Kant characterizes 'transcendental philosophy': "Transcendental philosophy is the act of consciousness whereby the subject becomes the originator [*Urheber*] of itself and, thereby, also of the whole object of technical-practical and moral-practical reason in one system" (21:78, 245). Being one's own maker establishes the basis for unifying theoretical cognition of natural mechanism (which Kant equates with the 'technical-practical' in 22:52) and practical cognition of human freedom under a common principle. This 'progression' from one to the other, from natural determination to freedom, is nominally a consequence of the OP's stated goal: establishing a progression from metaphysical principles of natural science to physics. Over the course of the work,

---

[10] See references to self-positing [*selbstsetzen*], or 'composition', throughout the entirety of OP, e.g. 22:326, 22:358, 22:384, 22:12.

however, it seems to constitute an ulterior goal altogether: "One must progress from subjective principles of appearance to what is objective in experience. One must progress from technical-practical to moral-practical reason, and from the subject as a natural being to the subject as a person" (22:49, 210).

1.2.2 Autonomy and reflectivity in self-positing ideas and maxims of reason

This ability to generate ('originate', 'give rise to', *urheben*) one's own thought, and thus the consciousness of one's own moral personhood, gives autonomy a new meaning: "Transcendental philosophy is autonomy, that is, a reason that determinately delineates its synthetic principles, scope, and limits, in a complete system" (21:59, 244). The concept of autonomy here, however, is broader than Kantian autonomy traditionally construed; here, autonomy characterizes Kant's conception of transcendental philosophy *in general* rather than a specifically moral principle. In fact, 'autonomy' now serves as the criterion governing the success of Kant's entire philosophical enterprise:[11] "Transcendental philosophy commences from the metaphysical foundations of natural science, and contains the *a priori* principles of the latter's transition to physics… without turning into heteronomy, it then progresses to physics" (21:59, 244).

This conception of autonomy cannot be dismissed as a mere turn of phrase; Kant reiterates it throughout the last fascicles of the manuscript,[12] particularly in describing the project of transcendental philosophy as such:

> Transcendental philosophy is the autonomy of ideas. (21:79)

> Transcendental philosophy is the subjective principle of ideas of objects of pure reason constituting themselves into a system, and of its autonomy according to the concepts: *ens summum, summa intelligentia, summum bonum*. (21:79, 246)

> This scale of ideas…. is *autonomia rationis purae*. (21:82, 248)

---

[11] One might situate this view not as a departure from Kant's critical position, but as a development of it. O'Neill (1990) has written convincingly on Kant's use of political metaphor in KrV, thus situating the project of critique as such as a distinctly *political* task. Moreover, in crucial passages of GMS, KrV, and KpV, Kant presents practical reason, and therefore autonomy, as also being evidenced by the cognitive subject's activity of self-consciousness (A546/B574, 4:457). However, as this paper aims to bring out, it is not until the post-critical period that this conception is developed into a broader account of autonomy than strictly *moral* autonomy.

[12] As Förster (1993) notes, the exact ordering of the manuscript is unknown: the pages, as ordered on Kant's desk when he died, were mixed together before they could be transcribed, and the manuscript suffered decades of neglect. Nevertheless, the precise ordering doesn't much matter for my purposes.

> The autonomy of ideas: to found experience as unity, *a priori*—not *from* experience, but *for* experience. (21:92, 253)

> *Ideas* are not mere concepts but laws of thought which the subject prescribes to itself. *Autonomy.* (21:93)

In these passages, autonomy is associated with the production of *ideas*. Such a connection may seem strange: to my knowledge, these passages constitute the only instances in Kant's texts in which autonomy is primarily linked with *thinking* ideas rather than *acting* in accordance with them. But here it may help to reflect on what ideas are. In KU, Kant characterizes them as follows: "Ideas, in the broadest sense, are presentations referred to an object according to a *certain principle* (subjective or objective) but are such that they can still never become cognition of an object" (5:342). That is, ideas are inherently norm-guided, insofar as "they are produced according to principles" (5:342). In both KrV and KU, Kant refers repeatedly to the 'maxims' that guide ideas of reason. I will return to the role of maxims in KU in the next chapter; for now, I will flag their role in KrV:

> I call *maxims* of reason all subjective principles that are obtained not from the character of the object, but from reason's interest concerning a certain possible perfection of the cognition of this object…. The systematic unity of the manifold knowledge of the understanding, as prescribed by reason, is a *logical* principle. Its function is to assist the understanding by means of ideas in those cases in which the understanding cannot by itself establish rules. (A648/B676)[13]

As this passage shows, already in KrV had Kant attributed a role to maxims in *theoretical* cognition in allowing cognizers to proceed in accordance with ideas that would otherwise not be open to cognitive representation. For example, Kant names several of the principles, or maxims, of natural teleology on which he focuses in KtU: "That the manifold respects in which individual things differ do not exclude identity of species, that the various species must be regarded merely as different determinations of a few genera, and these, in turn, of still higher genera… is a logical principle… without which there could be no employment of reason" (A651-2/B680-1). Reason requires subjective maxims because, in the absence of determinate rules, it must give itself its own normative course to follow in making sense of experience.

In KU, however, Kant reassigns the employment of such idea-guiding maxims from reason, in its *regulative* use, to judgment, in its *reflective* use. He writes, for example, that the "idea" of a natural purpose "is not a principle of reason for the understanding, but for the power of judgment… where, indeed, the judgment cannot be

---

[13] Thanks to Anja Jauernig for directing my attention to the Appendix to the Dialectic.

determining, but merely reflecting" (5:405). The ideas involved in thinking the systematicity of nature are no longer classified as regulative principles of *reason,* but are principles accorded to *judgment,* insofar as the configuration of our cognitive faculties requires them in order to organize experience into ordered knowledge.

1.2.3   Heautonomy

KU also introduces a new conception of autonomy specifically *for judgment.* In both introductions, Kant claims that the power of judgment "exhibits *autonomy*" insofar as it is "legislative with regard to the conditions of its reflection *a priori*" (20:225; compare 5:186). Kant specifies that the maxims that guide judgment are not "psychological" in origin, but normative: "They do not say what happens… but how they *ought* to be judged" (5:182). The normativity at issue makes judgment only *subjectively* autonomous, insofar as it legislates only to itself (rather than to all rational agents, through laws of freedom,[14] or to experience as such, through laws of nature). Thus judgment is not autonomous strictly speaking, but what Kant terms *heautonomous* (where the 'he-' prefix refers to the Greek reflexive pronoun, emphasizing its self-directed nature).[15]

That is, the power of judgment is subject to a self-given normative metaprinciple: in order to proceed in accordance with a given idea, it must methodologically dictate to itself the course of its own procedure, but this procedure can itself only be specified and ascertained in the course of the exercise of judgment itself. Judgment is thereby subject to an internal normative standard: its normativity is both subjective and made in the course of judgment itself. Calling judgment 'heautonomous', then, is another way of saying that the procedure of judgment is *self-made.*

The upshot of these shifts in Kant's views is that, on Kant's mature formulation, thinking ideas *requires* heautonomy, or the self-making of judgment, rather than reason under its regulative guise alone. Consequently, Kant's characterization of the subjective construction of a system of ideas as a form of 'autonomy', in the sense of self-making, is consistent with his views post-1790. The form of autonomy Kant has in mind, I posit, encompasses heautonomy. As I argue in this dissertation, moral

---

[14] For a convincing reappraisal of the heautonomy/autonomy distinction, see Kleingeld (2017).

[15] See Floyd (1998): "*Heautonomy* derives from the Greek definite article *he* being attached to the Greek pronoun for 'self', or 'itself': *auto.* The resulting term, *heauto,* means just what *auto* does except that it may only appear grammatically in a sentence reflexively (as in, e.g., 'I wash myself', 'he praised himself', etc.)… Philologically speaking, Kant is trying to emphasize a certain *necessarily* reflective character of the faculty of judgment" (205).

autonomy *already presupposes* heautonomy: moral actualization *also* requires the subjective, self-given generation of maxims. Thus, if Kant's characterization of reflective judgment in KU refers to cases that are *merely* [*bloß, nur*] reflective (rather than reflective *and* determining as in KrV,[16] or reflective *and* practical, as, I will argue, in GMS and KpV), 'autonomy', as the more general concept, will also comprise 'heautonomy'. Kant's frequent references in OP, not only to the making of (theoretical) ideas, but to 'moral-practical reason' and the generation of moral personhood, suggest that he had both conceptions—autonomy and the heautonomy it presupposes—in mind.

The 'self-making' of heautonomy constitutes a through line in the argument of OP, insofar as the conception of autonomy it describes is characterized in terms of an extensive process of self-making with both theoretical and practical dimensions. In theoretical terms, we 'make ourselves' by cognizing ourselves as objects: "Transcendental philosophy is the capacity of the self-determining subject to constitute itself as given in intuition…. *To make oneself*, as it were [*gleichsam sich selbst machen*]" (21:93, 254). In practical terms, we 'make ourselves' moral agents: "It is not even in the divine power to make a morally good man (to make him morally good): He must do it himself" (21:83). In both cases, the common principle being 'made' is personhood, both moral personality and the personhood of the rational cognizer:

> Transcendental philosophy is the doctrine of the complex of ideas, which contain the whole of synthetic a priori knowledge from concepts in a system both of theoretical-speculative and moral-practical reason, under a principle through which the thinking subject constitutes itself in idealism, not as thing but as person, and is itself the originator [*Urheber*] of this system of ideas. (21:91, 252)

Kant's repeated insistence in these passages on the fact that the idealist system comprises *both* 'theoretical-speculative' *and* 'moral-practical' reason, and that the 'constitution' of the subject requires *both* constituting oneself as a person (a moral notion) *and* 'originating' a system of ideas, provides additional evidence that the conception of autonomy at issue must encompass both theoretical *and* practical uses.

While this conception of autonomy is in tension with Kant's characterization of autonomy in GMS and KpV as limited only to the legislation of the *moral law,* Kant's characterization of heautonomy shows the contours of an account that cuts across the theoretical-practical distinction.[17] Kant characterizes the heautonomous principles of judgment in the Second Introduction of KU as "themselves fit neither for theoretical

---

[16] See discussion in next chapter.

[17] This possibility is already latent in Kant's characterization of ideas of reason in KrV: "Perhaps the transcendental ideas of reason make possible a transition from the concepts of nature to the practical concepts and in this way provide for the moral ideas themselves support and coherence with reason's speculative cognitions" (A329/B386).

nor for practical use", thus "mak[ing] possible a transition" from one domain to the other (5:176). In the First Introduction, Kant specifies that the use of judgment is 'technical', or practical in the contemporary sense (rather than practical as moral): "The reflecting power of judgment thus proceeds… not schematically, but *technically,* not as it were merely mechanically, like an instrument, but *artfully* [künstlich]"[18] (20:213-4). Reflective judgment is practical in the sense that an artist is: proceeding freely and creatively in action and thought with the principles given to oneself, rather than undertaking actions of directly moral import.

Moreover, Kant's repeated references throughout OP to 'organic matter', 'organisms', and 'organic bodies' (there are 140 instances of 'organic' alone) bolster the suggestion that reflective (specifically teleological) judgment becomes increasingly central to the expansive conception of autonomy he now endorses. Among these references is also an explicit reference to the 'autonomy' presupposed by our concept of an organism, suggesting that Kant really had broadened his conception of autonomy to include even straightforward instances of heautonomy: "Note. Of the autonomy of the concept of the organization of matter, without which we ourselves would have no organs" (22:86, 191). What Kant has in mind with his conception of 'self-making' also involves this consciousness of ourselves as purposive, corporeal beings: "Even the organism is contained in the consciousness of oneself. The subject makes its own form in accordance with *a priori* purposes" (22:78, 186). We 'make ourselves' not just by constructing systems of ideas (including the idea of one's self as a human being), but as embodied, purposive organisms; the purposive configuration of our bodily being is, in part, up to us (for example, through the development of physical skills and capacities).[19] Both of these notions, however—proceeding in accordance with the maxims necessitated by ideas, as well as conceiving of ourselves as purposive organisms, as I explain below—require the use of reflective judgment.

1.3   Self-making and heautonomy in the *Doctrine of Virtue*

Against the application presupposition, Kleingeld rightfully observes that the notion of self-legislation becomes increasingly less relevant for Kant's mature philosophy, since this notion also lacks a central place in the OP. What poses a problem for her

---

[18] This term is often also translated as 'artistic', although 'artificial' is its much more prevalent contemporary meaning. Another possible translation is 'manmade', which would emphasize the connection to the self-making of judgment. Kant's main point here is that reflective judgment establishes its own principles rather than having them be externally dictated.

[19] The connection between self-making and purposiveness suggests a Kantian reference point for a position most commonly ascribed to Aristotle; see Korsgaard (2009) on the Aristotelian elements of her account of self-constitution.

account, however, is the fact that Kant nevertheless *retains* the concept of autonomy in OP, even without the analogy to legislation she takes to be definitional. Indeed, autonomy in OP is no longer governed by the metaphor of self-legislation (giving a *law* to oneself), but that of self-making (*making* oneself). Thus, autonomy does not drop out of Kant's mature philosophical position, but it does undergo an important shift.[20] As if to emphasize the distinction between the critical and post-critical conceptions of autonomy, Kant populates the latter notion with a corresponding new constellation of terms: self-constitution [*sich constituieren*], self-creation [*sich schöpfern*], self-origination or self-authorship [*sich urheben*], self-construction [*sich zimmern*]. While the notion of self-legislation is *juridical*, the notion of self-making is more explicitly *generative*. The subject as self-maker is no longer conceived on the model of an ideal moral judge deliberating over an already given law, but instead on that of creating a system of (ultimately self-directed) ideas. Autonomy here has to do with 'founding' oneself rather than legislating oneself, with self-production rather than self-constraint: "According to transcendental idealism, the subject constitutes itself… as a being who is founder [*Begründer*] and originator [*Urheber*] of his own self" (21:14).

But is this conception of autonomy relevant to Kant's practical philosophy? On the virtually universal application presupposition, the answer would be a resounding 'no'. On this presupposition, the definition of autonomy presented in GMS is simply 'applied' to Kant's later moral writings, even if this conception no longer appears in these writings. Thus, as this near-universal consensus among scholars would have it, the conception of moral autonomy in GMS is much more relevant to texts such as MS than the not specifically moral conception presented in OP. After all, the title of the former work makes clear that it is to lay the *groundwork* for Kant's metaphysics of morals, whereas many scholars have dismissed OP as the product of Kant's senility.

Against this common assumption, there is textual evidence *within MS itself* that Kant took the OP to provide a *model* for a metaphysics of morals. In DV, Kant explicitly draws a parallel between the aim of this work and OP, suggesting that both texts, composed in overlapping time periods in the mid-1790s, should be treated as complementary:

> Just as a passage from the metaphysics of nature to physics is needed—a transition having its own special rules—something similar is rightly required from the

---

[20] While Kleingeld makes no reference to OP, scholarship specifically on OP has also missed this point. In this literature, it is generally accepted that the OP presents a new account of theoretical reason; however, the conception of practical reason Kant advances therein has either been neglected, or read as conforming unproblematically with Kant's critical moral philosophy (see, e.g., Förster 1993). Yet such posing of an equivalence between the two elides the distinctions between the two accounts.

> metaphysics of morals: a transition which, by applying the pure principles of duty to cases of experience, would *schematize* these principles, as it were, and present them as ready for moral-practical use. (6:468)[21]

That is, the aim of OP is to secure a transition from the *a priori* to the empirical in the domain of theoretical philosophy, while MS requires a similar transition in the moral domain.

In noting this important point of commonality between the two works, I do not want to be taken to conflate their differences. Indeed, OP and MS could be seen to operate on two distinct intellectual registers: OP, at least at the outset, seems to be situated transcendentally, proceeding from its 'first principle' of self-positing; MS, if read as proceeding from the transcendental 'foundation' secured by GMS, could be taken to be merely supplementary or additive, and thus not transcendental in its own right. However, the incongruity between them appears less straightforward on closer inspection: OP aims to secure a transition from the transcendental to the empirical, and therefore cannot remain solely transcendental, while MS incorporates both *a priori* metaphysical first principles and a 'practical anthropology'—the latter of which Kant newly acknowledges as a necessary component of his moral system.[22] Thus, while there are important methodological differences between the two works, both aim at achieving *cohesion* between metaphysical principles and the empirical conditions to which they apply. Arguably, this need for both domains to *cohere* is a more pressing concern at this stage than in Kant's earlier works and, as I suggest below, may have been the direct result of the conclusions reached in KU.

The connection Kant draws between the respective functions of OP and MS

---

[21] Some might take the 'metaphysics of nature' in this passage to refer to KrV rather than the *Metaphysical Foundations of Natural Science* (see Friedman 1998: 168); on this reading, the 'passage' from the metaphysics of nature to physics would be the *Metaphysical Foundations* rather than the OP. However, in the A-edition of KrV, Kant differentiates the project of this text from a 'metaphysics of nature': "Such a system of pure (speculative) reason I myself hope to provide under the title of *Metaphysics of Nature*. That system, though not half as voluminous as this critique, is to be incomparably richer in content. But first the critique had to establish the sources and conditions of the possibility of that system, and needed to clear and level a ground that was entirely overgrown" (Axxi). In the *Metaphysical Foundations* itself, Kant seems to take the subject matter of this work to be the metaphysics of nature (4:470). Yet in KU, Kant reiterates that the metaphysics of nature is a separate project yet to be completed, and which would complement the metaphysics of morals—seemingly reinforcing the need for a transition and the inadequacy of the *Metaphysical Foundations* to fill this role on its own (5:170). It seems clear, then, that the 'passage' from metaphysics of nature to physics cannot refer to the *Metaphysical Foundations* (as the passage from KrV to physics), but must refer to the transition project of OP. See discussion of the shifting referent of 'metaphysics of nature' in Pluhar (Bxliiin, 40), Guyer and Wood (1998: 66), and Friedman (1998: 243, 260).

[22] I return to the role of anthropology below.

suggests that the reappearance of autonomy is not attributable merely to the revision of his views on *political* legislation (as Kleingeld alleges), nor to Kant's senility in OP, but to a more pervasive problem posed by Kant's late practical philosophy:[23] namely, that the notion of self-legislation alone is insufficient to account for the connection between pure principles and contingent circumstances, which instead requires an independent contribution by the moral agent. Thus, the need for a transition is felt not only from metaphysics to natural science, but also from pure moral principles to situated practical deliberation.

In OP, it is autonomy, conceived in terms of self-making, that secures the transition from pure principles to empirical experience. In MS, it is judgment, through casuistical applications of pure moral principles to contingent cases, that helps to secure this transition in the moral sphere (see diagram):

	Transition from ⇨ to:	Transition from ⇨ to:
Pure	Metaphysical foundations	Universal moral maxims
⬇	**Opus Postumum** (autonomy as self-making)	**Metaphysics of Morals** (practical judgment)
Empirical	Natural science	Concrete cases

As we saw, however, Kant's characterization of autonomy in OP, as the construction of ideas into a system, already presupposes the autonomy of *judgment,* insofar as the maxims required for the regulative employment of ideas are reflexively given to judgment, constituting its heautonomy or self-referential normativity. Thus, autonomy as self-making in OP also requires the role of judgment. But can this connection shed any light on *practical* judgment, the point of similarity Kant emphasizes between MS and OP?

1.3.1   From the denial of a practical schematism to a 'reflective faith'

---

[23] I don't want to be taken to suggest that this is the *only* problem Kant's new conception of autonomy is intended to solve. The more pervasive problem, I take it, is that of finding a fundamental principle that can unify the spheres of theoretical and practical cognition. Kant felt this to be a problem well before OP; see Neuhouser (1990: ch. 1) on Kant's search for a deduction of the categorical imperative to establish a structural affinity to the transcendental deduction of the categories. However, it seems to have become more pressing in his final decade; see Kant's correspondence with Maimon (11:285) and Weiss (12:185-6), who press Kant on his 'universally valid principle' and 'actual ground' of his critical philosophy, respectively.

Kant continues the passage in DV in which he draws this explicit parallel between the 'schematizing' functions of OP and MS by suggesting that, in the moral case, a transition is required in order to answer questions such as:

> How should one behave, for example, toward human beings who are in a state of moral purity or depravity? Toward the cultivated or the crude? Toward the learned or the unschooled, and toward the learned in so far as they use their science as members of polite society or outside society, as specialists in their field (scholars)? Toward those whose learning is pragmatic or those in whom it proceeds more from spirit and taste? How should people be treated in accordance with their differences in rank, age, sex, health, prosperity, and poverty, and so forth? These questions do not yield so many different *kinds* of ethical *obligation* (for there is only *one*, that of virtue as such) but only so many different ways of *applying* it. (6:468-9)

As Kant argues throughout DV, ethics involves 'imperfect duties' that necessitate a certain 'latitude' in how to fulfill them.[24] Thus ethics "unavoidably leads to questions that call upon judgment to decide how a maxim is to be applied in particular cases" (6:411). That is, to answer questions such as the ones Kant raises in this passage—how to treat different *kinds* of individuals, in accordance with their social roles, capacities, and contingent characteristics—we can't rely on explicit *rules* directing us what to do in each case (since the variety of human subjectivity is immeasurable), but must use our individual[25] capacity of judgment. This is because the role of practical judgment, if conceived as mechanically applying rules, would give rise to a regress, one akin to the regress Kant notes in his discussion of theoretical judgment in the Analytic of Principles of KrV: the application of each rule would need to be guided by the introduction of a new rule, and so on ad infinitum (A133/B172). The role of judgment therefore does not admit of systematization. Kant concedes that here "ethics falls into a casuistry", which can only "fragmentarily", not systematically, be incorporated into ethics (6:411). Thus, casuistry cannot be conceived as a "science", but only as a "practice", specifically, the practice of moral judgment (6:411). Because matters of judgment,

---

[24] Though, as I will argue in Chapters 3 and 5, perfect duties are also not reducible to mere *applications* of principle to experience, but are also subject to 'moral conceivability', or to how they are conceived or described. In the case Anscombe (1956) raises of Truman dropping the atomic bomb on Japan, the apparent moral dilemma arises in part due to collective uncertainty about how wide a scope to attribute to 'killing', as well as the moral conflict at stake in potentially saving millions of lives (though Anscombe would vehemently disagree with both sources of uncertainty). Thanks to Don Garrett for pressing me for this example.

[25] Or collective—some of these questions are also subject to collective deliberation. In either case, the determination of what to do remains perpetually open to negotiation rather than being laid down as a rule once and for all. I return to this point in later chapters.

such as how to treat individuals in light of their varied conditions and social roles, therefore resist complete systematization, Kant claims that casuistical applications of ethical obligation, though necessary for a "complete presentation of the [moral] system", can only be "appended" to it fragmentarily (6:469).

This view of practical judgment explicitly contrasts with the position Kant advances in KpV (1788), where he rejects the role of a schematism for practical reason, affirming only the possibility of introducing further rules. In the Typic of Pure Practical Judgment, Kant writes that since judgment applies universal principles to particular cases, and since all particular cases can only be empirical, "it seems absurd to want to find in the sensible world a case… to which there could be applied the supersensible idea of the morally good" (5:68). Thus, practical judgment is subject to "special difficulties" that theoretical judgment does not face, since it would seem impossible to apply laws of freedom to events of the sensible world (5:68). Consequently, Kant here denies that practical judgment, unlike theoretical judgment, could constitute a 'schema': "No intuition can be put under the law of freedom… and hence no schema on behalf of its application *in concreto*" (5:69). In doing so, Kant *rejects the very point of connection he draws between OP and MS above*—namely, that *both* theoretical *and* practical reason require a 'schematism' between universal principles and conditions of application, and with it an *irreducible* role for practical judgment.[26] Because, at this earlier stage, Kant holds that practical reason does not admit of a schematism, he admits only of the possibility that practical judgment take the form of a *law* (5:69). Thus, he explicitly rules out the characterization of judgment given in KrV and MS (thus both theoretical *and* practical) as a 'practice' which, due to the threat of a regress, precisely *cannot* be formalized as a law.

How should we make sense of this significant change in Kant's conception of practical judgment? I suggest that the discovery of reflective judgment in KU was more broadly transformative for Kant's mature practical philosophy than has generally been recognized. As Makkreel (2002) and Munzel (1999) have noted, the role of 'reflection' is explicitly mentioned in the *Religion* (1794) as what is required to effect a transition to a moral community.[27] Kant refers there to a moral faith which is "*reflective*

---

[26] To recall, the complete passage is the following: "Just as a passage from the metaphysics of nature to physics is needed—a transition having its own special rules—something similar is rightly required from the metaphysics of morals: a transition which, by applying the pure principles of duty to cases of experience, would *schematize* these principles, as it were, and present them as ready for moral-practical use" (6:468).

[27] The specific passage they cite is: "The basis for the transition to the new order of things… once grasped after mature reflection [*Überlegung*], will be carried to effect" (6:122). Since Kant here uses 'Überlegung' rather than 'Reflexion' (the term employed in KU to refer to reflective

[*reflectirenden*]" (6:52), insofar as it allows reason to "extend itself to extravagant ideas," which it "cannot incorporate… into its maxims of thought and action" (6:52). Moreover, this 'reflective' mode of faith is required if "something more" from the "inscrutable field of the supernatural" can be brought to the understanding, since "reason even counts on this something being made available to its good will even if uncognized" (6:52).[28] That is, Kant attributes a specifically *moral* role here to *ideas*, which are present to a good will even if they cannot be cognized, or integrated into maxims or (moral) concepts. Ideas—concepts beyond our cognitive grasp, that elide formulation in specific maxims of action—nevertheless have an important role to play in *guiding* our formation of maxims (our 'good will'). Moreover, it is only through *reflective judgment* (a 'reflective faith') that such ideas can make themselves available to us in this morally constructive fashion.

If Kant affirms the role here of reflectivity in practical reason, this suggests that heautonomy plays an analogous role in moral deliberation to theoretical reason. While Kant does not, to my knowledge, explicitly refer to reflective judgment in MS or elsewhere in the *Religion*, Kant's discovery of reflective judgment would clarify key changes in Kant's account of practical judgment from GMS to MS: the independently necessary role of practical judgment, the necessity that judgment be *practiced* rather than governed by rules, and judgment's necessarily 'fragmentary' status in his philosophical system. In each of these respects, Kant's account of practical judgment comes to much more closely resemble his account of judgment in KrV, suggesting that he comes to reject the disanalogy initially established in KpV between theoretical and practical judgment (and with it, affirming the possibility of a practical schematism on which the structural affinity between OP and MS depends).

If so, many of Kant's remarks in MS can be taken to refer to the *reflective* dimension of practical reason, in addition to the *practical* dimension as it has been commonly understood. For example, casuistry, insofar as it involves the subsumption of particular cases under universal principles in the service of forming a complete ethical system,

---

judgment), which is closer to 'consideration' than 'reflection' in Kant's sense, this passage provides weaker textual support than the passage I cite here.

[28] The full passage is the following: "Reason, conscious of its impotence to satisfy its moral needs, extends itself to extravagant ideas which might make up for this lack, though it is not suited to this enlarged domain. Reason does not contest the possibility or actuality of the objects of these ideas; it just cannot incorporate them into its maxims of thought and action. And if in the inscrutable field of the supernatural there is something more than it can bring to its understanding, which may however even be necessary to make up for its moral impotence, reason even counts on this something being made available to its good will even if uncognized, with a faith which (with respect to the possibility of this something) we might call *reflective*" (6:52)

may be taken to constitute a form of reflective judgment. As Kant presents the 'transition' requisite for a metaphysics of morals in this passage, the formulation of specific principles as to how to treat individuals who are 'cultivated or crude', 'morally pure or depraved', or of different races, sexes, or classes must not be undertaken *a priori,* but can only result from having compared and considered different particular cases. Thus, the form of judgment necessary to secure the relevant schematism accords with Kant's definition of reflective judgment in the First Introduction to KU, where he claims that "to reflect… is to compare and to hold together given representations either with others or with one's faculty of cognition, in relation to a concept thereby made possible" (20:211). Reflective judgment *incorporates* particular cases into the formation or further determination of concepts, rather than merely subsuming them under pregiven concepts. Casuistry, if understood as the formulation of ethical principles which specify how to proceed in determinate cases, would have to be conceived as a form of reflective judgment, as well as practical judgment.

This outcome can be taken to accord with Kant's shift in emphasis in his mature moral philosophy. By the 1790s, Kant comes to emphasize the cultivation of virtue as what enables human moral actualization over the metaphysical status of morality as such. Thus, as Guyer (2013) has argued, autonomy must be conceived not just as the rational capacity accessible to all moral agents, Kant's focus in GMS and in the 1780s more generally; instead, by the 1790s, Kant stresses that autonomy must also be empirically realized in the lives of given agents—and on this empirical level, one's 'susceptibility' to being moved by duty requires cultivation.

Perhaps, then, Kant's late moral philosophy was transformed in ways the application presupposition has obscured from view. For example, if, as Herman notes, "there is the surprising fact that the CI [categorical imperative] procedure is not used in either part of the *Metaphysics of Morals*", this suggests the possibility that Kant's considered conception of moral philosophy *came to deemphasize* the role of the categorical imperative procedure given pride of place in GMS (1993: 133). This is not to say that the categorical imperative becomes irrelevant; Kant continues to refer to it (though less frequently) in both texts. However, Kant's account of practical deliberation appears increasingly to *broaden beyond* mere applications of rules, including the categorical imperative test. The picture we get of practical deliberation in OP and MS seems to put more emphasis on the individual's role in *schematizing* moral ideas (including virtue), requiring their individual practice of judgment.

1.3.2 *Schematizing the anthropological and the moral*

In doing so, a more immediate point of connection is established between Kantian anthropology and Kantian morality. Initially merely a domain of theoretical investigation of the empirical human being (akin to psychology), Kant comes, throughout the 1780s, to characterize it as '*practical* anthropology', and thus as a requisite domain of moral philosophy.[29] In ApH, Kant claims that anthropology, now in its practical or 'pragmatic' sense, concerns "what *he* [the human being] as a free-acting being *makes of himself*, or can and should make of himself" (7:119, my emphasis).[30] That is, anthropology, if understood practically, *also* centrally concerns a kind of 'self-making'.[31]

Moreover, anthropology in this sense is continuous with Kant's account of virtue, since it is not just descriptive ('what the human being makes of himself'), but *normative* (what the human being 'can and should make of himself').[32] The aim of anthropology is also to depict possibilities for the human subject's course of moral actualization. It is both empirical—describing how human beings actually *do* actualize or 'make' themselves as subjects—and practical—advocating for their *moral* self-actualization or 'self-making'.[33]

In MS, Kant draws another structural parallel between this text and OP. If OP is to provide the "step which connects both banks" (21:403, 15) between empirical natural science and a metaphysics of nature, MS holds the task of integrating empirical science of human nature (anthropology) into a metaphysics of morals:

> Just as there must be principles in a metaphysics of nature for applying those highest universal principles of a nature in general to objects of experience, a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular *nature* of human beings, which is cognized only by experience, in order to *show* in it what can be inferred from universal moral

---

[29] See VA Mrongovius (1784-1785) 25:1367 for Kant's first reference to 'practical anthropology'. See also discussion in Louden (2012), as well as Wood (1991, 1999).

[30] As distinct from anthropology in its 'physiological' sense, as Kant had first characterized it in his earliest lectures, which "concerns the investigation of what *nature* makes of the human being" (7:119).

[31] For an analysis of ApH's account of 'self-making', see Wood (1999, 2003).

[32] The normative dimension of Kant's considered view of anthropology has often been missed; for a recent example, see Theunissen (2016)'s characterization of anthropology as purely descriptive and consequently "not strictly speaking a branch of moral philosophy at all" (109). As I argue, there is textual evidence against both claims, and thus against Theunissen's conclusion—that a metaphysics of morals cannot contain any empirical considerations—as well. Theunissen's interpretation relies heavily on GMS and KpV, not taking into consideration the changes in Kant's views between the publication of these texts and MS proper I have noted here.

[33] I return to Kant's characterization of virtue in Chapter 2 below, of which I advance a reading in terms of heautonomy.

principles. (6:217)

If a metaphysics of morals will often involve taking as its object 'the particular nature of human beings', this refers to the domain Kant marks out for *anthropology*. At the same time, Kant increasingly comes to conceive of the domain of anthropology as having a distinctly *practical* component (25:1367, 27:1398). Thus, Kant draws together the two domains such that each comes to include aspects of the other. A metaphysics of morals has the task of schematizing universal *a priori* principles and conditions of application, which in practice will also require schematizing our self-understanding as at once anthropological, corporeal, socially situated beings and moral agents.[34] As Kant emphasizes here, the empirical realization of autonomy depends not just on cultivating one's 'susceptibility' to moral duty (Guyer 2013), but also reflection on how universal *a priori* moral principle should be actualized in contingent circumstances which, often, are not immediately up to us.

1.4   Reflective judgment and practical reason

Given what I have said so far, the role of reflection in moral deliberation may look like it had no roots in Kant's initial views. Yet, in certain respects, reflective judgment had long been basic to Kant's account of practical reason, putting further pressure on the application presupposition. In the remainder of this chapter, I will take up two examples of the role played by natural teleology, one of the two main results of reflective judgment in KU, from GMS to MS: the formula of the realm of ends (FRE) and self-perfection as a moral notion. I take each case to demonstrate the longstanding role of reflective judgment to Kant's account of practical reason. However, I take the shift in Kant's appraisal of self-perfection, in particular, to show that the discovery of reflective judgment allowed Kant to broaden his conception of practical reason to accommodate the *reflective dimension* of independently formulating subjective maxims of action.

1.4.1   *Natural teleology in deriving FRE*

---

[34] See KU 5:435 for Kant's claim that in order to connect our conception of nature to the human being as a moral subject, we require the thought of the 'culture' of the empirical human being, including anthropological considerations such as human skill, civil society, and the effects on humanity of war, luxury, and education (5:431-3). Zammito (1992: 267) and Allison (2001: 205) both position the task of schematizing the empirically conditioned human being and the autonomous moral agent as the principal 'gulf' KU is attempting to bridge.

In GMS, Kant's derivation of the formula of the realm of ends, FRE[35], directly results from a comparison to natural teleology, later accorded to reflective judgment in KU. Kant claims, "All maxims from one's own legislation ought to harmonize into a possible kingdom of ends as a kingdom of nature" (4:436). In a footnote, Kant reinforces the parallel to natural teleology:

> Teleology considers nature as a kingdom of ends, moral science a possible kingdom of ends as a kingdom of nature. There the kingdom of ends is a theoretical idea for explaining what exists. Here it is a practical idea for the sake of bringing about—in conformity with precisely this idea—what does not exist but can become actual by means of our behavior. (4:436n)

That is, formulating the maxim of the kingdom of ends requires a *dis*analogy to natural teleology: it requires, first, the function of reflective judgment, specifically *teleological* judgment, to think the concept of natural purposiveness (as a 'kingdom of nature'); and second, this concept serves as a point of comparison or contrast, a *disanalogy*—which must also be formulated by reflective judgment—on which the moral maxim, that of the kingdom of ends, depends. Kant reiterates the same point immediately before introducing the formulation of this maxim as the "supreme principle of autonomy" (4:440): "A kingdom of ends is possible only according to the analogy with a kingdom of nature—but the former just according to maxims, i.e. self-imposed rules, the latter just according to laws of externally necessitated efficient causes" (4:438). Since the 'kingdom of nature' as Kant characterizes it in KU[36] requires the thought of nature as an *externally purposive* system, formulating the maxim of autonomy requires thinking of other rational beings as related to each other in comparison to such a system. That is, in the domain of nature, each organism is thought as serving a determinate purpose in a natural chain, as an 'externally necessitated efficient cause'—the grass nourishes the deer, the deer nourishes the wolf, and so on. By contrast, in the domain of practical reason, each rational being must be thought as interrelated with all others *not* to serve a determinate, preestablished purpose for someone else,[37] but in the sense that their end is up to them to determine—but only if this end is one that could be

---

[35] The maxim is given several paragraphs later: "Act according to the maxims of a member universally legislating for a merely possible kingdom of ends" (4:439)

[36] See Kant's references to various 'kingdoms of nature' (5:426-7).

[37] As Kant establishes in KU, the human being is the *only* natural entity whose purpose, as a link in the chain of nature, seems unclear; indeed, it is this lack of necessity of the human subject leads to the thought of the *moral* subject as the final purpose of nature: "Yet one does not see why it is necessary that human beings exist… thus one does not arrive at any categorical end, but all of this purposive relation rests on a condition that is always to be found further on, and which, as unconditioned, (the existence of a thing as a final end) lies entirely outside of the physical-teleological way of considering the world" (5:378).

determined, or accepted, by everyone else. If FRE states that I must be able to legislate an end, or purpose, that every other rational being can accept (precisely *because* our purposes, as rational beings, are ones up to us to determine), the analogy to natural external purposiveness proves to be a *dis*analogy. The ends of a rational being can, for the most part, only be constrained negatively rather than determined positively, since they can be ruled out if they fail to be universalizable, but cannot be imposed from without if they are to count as free.[38] However, Kant seems to think that the principle of the external purposiveness of nature is required to fully make sense of this contrast between rational and natural beings, and thus to vindicate FRE.[39]

1.4.2   Natural teleology in Kant's shift on self-perfection

In MS, the connection between natural teleology and practical reason is even more robust. As Kant emphasizes in DV, the conception of self-perfection appropriate to moral philosophy is not *transcendental,* but *teleological.*[40] Understood 'teleologically'—through reflective judgment—perfection refers to "qualitative perfection", or "the harmony of a thing's properties with an *end*" (6:386). Only the teleological conception of perfection admits of more than one kind, since the transcendental conception "can only be one" (6:386). However, moral self-perfection can only be conceived as the former—on the teleological conception. In order to conceive of oneself as the potential object of self-improvement, one must conceive of oneself as the end of a teleological progression: at once an empirical, anthropological human subject and the bearer of the

---

[38] As O'Neill (1975) discusses at length, Kant does think that there are certain 'special' ends which are duties and thus positively determined, such as self-perfection and the happiness of others.

[39] Although a systematic account of this is outside of my scope, such an analysis of natural teleology in GMS could be extended further. For example, in Section I of GMS Kant motivates the notion that the "true function" of practical reason is to produce a good will by appealing yet again to natural purposiveness: "In the natural predispositions of an organized being, i.e. one arranged purposively for life, we assume as a principle that no organ will be found in it for any end that is not also the most fitting for it and the most suitable" (4:395). Many other such examples populate GMS and KpV.

[40] The complete passage is the following: "The word *perfection* is open to a good deal of misinterpretation. Perfection is sometimes understood as a concept belonging to transcendental philosophy, the concept of the *totality* of the manifold which, taken together, constitutes a thing. Then again, as a concept belonging to *teleology*, it is taken to mean the harmony of a thing's properties with an *end*. Perfection in the first sense could be called quantitative (material) perfection, and in the second, qualitative (formal) perfection. The quantitative perfection of a thing can be only one (for the totality of what belongs to a thing is one). But one thing can have several qualitative perfections, and it is really qualitative perfection that is under discussion here." (6:386)

ideal of moral perfection as one's ultimate end, which is the product in turn of one's ever-present capacity to act in accordance with moral duty.

The distinction Kant draws here between transcendental and teleological conceptions of perfection suggests that Kant's mature elaboration of the cultivation of virtue through self-perfection is contingent on his account of reflective judgment in KU; in an important respect, Kant was unable to formulate the former prior to the latter. Indeed, in GMS Kant only acknowledges the possibility of the "ontological concept" of perfection, which Kant rejects as heteronomous (4:443). In KU, Kant defines this ontological concept as "transcendental perfection", contrasting it, as he does once more in MS, with teleological perfection (20:228). While the former can only be singular, admitting of the teleological conception of perfection allows us to "speak of *a* perfection (of which there can be many in a thing under the same concept of it)", in which case the type of perfection at issue is "grounded in the concept of something, as an end" (20:228). Thus, in GMS Kant appears not to have yet admitted the possibility that self-perfection may take a different form than the transcendental, and thus that teleological judgment may be needed to complete his picture of virtue.

In KpV, Kant newly differentiates between "theoretical" (or "transcendental" and "metaphysical" concepts of perfection) and "practical" perfection, which accords with the teleological definition of perfection as "the fitness or adequacy of a thing for all sorts of ends"; considered "as a characteristic of the human being… is nothing other than talent and… skill" (5:41). While Kant once again rejects the moral role of perfection (including in its 'practical' or teleological sense), he does so insofar as it is taken to constitute a practical determining ground of the will. That is, taking perfection to be the sole principle of morality, as Christian Wolff does, fails to confer a procedure on which subjects can determine for themselves which ends of action are appropriate; instead, it requires that ends "must first be given to us", and thus constitutes heteronomy (5:41).[41]

As I read the progression from GMS to KpV to MS, in KpV Kant first introduces a distinction between transcendental and teleological modes of perfection, associating the latter with practical self-perfection. But, in KpV, he still lacks the distinction between cultivating skill and cultivating discipline, or between the general aptitude for pursuing various ends and the ability to choose which ends are appropriate (that is, moral) ones to act on. The skill/discipline distinction is what Kant will elaborate in

---

[41] On this point, see Guyer (2004) on Kant's differences with Wolff when it comes to self-perfection. As Guyer notes, Kant initially rejects perfectionism; however, the version he comes to endorse puts emphasis on perfecting one's powers of *choice* rather than perfecting one's *condition* (312).

§83 of KU. The characterization of self-perfection Kant advances in MS is contingent on the "negative" notion of discipline, or the "training" of one's will, rather than skill (5:431). In MS Kant refers to self-perfection as the "cultivation of one's *will* (moral cast of mind), so as to satisfy all the requirements of duty", as well as the cultivation of "those concepts that have to do with duty" (6:387).

Consequently, in MS Kant advances a conception of self-perfection that newly retains the connection to moral duty he had first held that this notion was lacking. But arriving at the appropriate formulation of this notion required the account of teleological judgment, and with it reflection on self-cultivation (as discipline rather than skill), Kant first elaborated in KU.[42] In KU, it is the thought of the internal perfection of natural organisms that vindicates the thought that they are internally purposive beings; the outcome presented in this course of activity of reflective judgment is *teleological* perfection, the thought of a perfection of a thing in relation to a possible end or purpose (5:374). In MS, the appropriate (teleological) conception of self-perfection results only from a comparison between the internally purposive character of natural entities and our own potentially moral purposive action as rational agents. We share similar cellular processes and modes of regeneration as other natural organisms; we, too, are internally purposive organisms. Yet we are also *rational* beings, with the ability to *perfect ourselves* in the direction of our own choosing. Thus, while in both KU and MS Kant stresses that teleological perfection allows for the possibility of many possible kinds of perfection under the same concept, this point really only begins to attain force when it comes to our self-conception as perfectible and potentially moral beings. Kant claims repeatedly in MS that every individual must follow her own self-given procedure in perfecting herself, and that this involves the individual derivation of one's own maxims (concepts of duty): "It is a contradiction for me to make another's *perfection* my end…. For the *perfection* of another man, as a person, consists just in this: that he *himself* is able to set his end in accordance with his own concepts of duty" (6:386). One particular agent's chosen procedure of self-perfection may not be valid for another. Thus, while moral self-perfection ultimately aims at the ostensibly "*narrow* obligation" of doing one's "duty *from duty*", Kant insists that the law does not prescribe any specific "action" that would fulfill this duty; the moral agent must set her own individual course in determining which actions would best fulfill this end (6:392).[43]

---

[42] On Kant's rejection of perfectionism, see Rawls (2000: 228); on Kant's eventual embrace of a conception of self-perfection in terms of choice of ends, see Guyer (2011).

[43] Although Kant does have certain broad suggestions; see, e.g., Kant's discussion of 'powers of the soul' at 6:445 or the cultivation of the feelings of beneficence, gratitude, and sympathy at 6:452-8.

The course of reflective judgment therefore extends *beyond* that delineated in the *Critique of Teleological Judgment:* the same intermediate conclusions vindicated there in the course of the procedure of teleological judgment—external and internal purposiveness—then serve as premises in the derivation of new moral maxims in GMS and MS, namely FRE and self-perfection. Where FRE, as the maxim of autonomy, is vindicated by its dissimilarity to *external* purposiveness, the appropriate conception of self-perfection is vindicated by its dissimilarity to *internal* purposiveness. Autonomy requires us to consider the *context* of our moral action: namely, the abstracted community of other rational beings, who must be conceived in contrast with the *external* purposiveness observable in nature. Self-perfection, meanwhile, requires us to reflect reflexively on our *own* nature as one intrinsically dissimilar from *internal* purposiveness.[44] In both cases, the purposiveness at issue is one that we can choose (or self-legislate) rather than one that is externally imposed; the analogy to nature enables us to reflect on our *dis*similarity with other natural organisms. In both examples, reflective judgment is required in order for the requisite *a priori* maxims to be properly conceived and derived.

As I will argue in the next chapter, however, reflective judgment is also involved in the formulation of subjective, subsidiary (or 'empirical') maxims. Indeed, in both FRE and self-perfection, the point of dissimilarity to natural organisms hinges on what I characterize as the heautonomy of practical judgment in human agents: the role played by our free choice to formulate and revise our maxims of action (that is, to choose the ends on which to act). In the case of self-perfection, for example, Kant suggests that reflective judgment is required not just to conceive of the appropriate definition of self-perfection in the abstract, but also to undertake a specific *course* of self-perfection as a given moral agent. For example, Kant articulates the command of moral self-perfection as "'be perfect.' But a human being's striving after this end always remains only a progress from *one* perfection to another" (6:446). That is, the course of realizing the end of a given mode of self-perfection results in the adoption of a *new* end of self-perfection; the formulation of new ethical ends (or rules) is achieved only in *practice,* in the

---

[44] Both cases of comparison are *disanalogies* of a sort. As I said above, it is really in *contrast* to external purposiveness that FRE becomes intelligible. Moreover, while Kant defines 'teleological perfection' similarly in KU and MS, the notion of one entity's having 'many possible perfections' only seems fully applicable to the case of a rational agent who can choose a particular direction of self-perfection. In both cases, the analogy to nature serves to put into relief our unique capacity to freely determine our ends. However, a productive contrast or disanalogy serving to better elucidate or define a concept, principle, or rule is an instance of reflective judgment; see Kant's discussion of 'difference' as a defining function of reflection in the Amphiboly (A260/B316).

course of undertaking a self-given procedure. Relatedly, undertaking one's own self-perfection must, in a certain respect, be not just autonomous, but also *heautonomous*: one's own procedure of self-improvement must accord with a normativity one claims for oneself, since this procedure is valid only for the particular agent who chooses to adopt it (even as it is *also* governed by universal general principles). Thus, in the case of natural or pragmatic self-perfection, Kant notes,

> Which of these natural perfections should take *precedence*, and in what proportion one against the other it may be a man's duty to himself to make these natural perfections his end, are matters left for him to choose in accordance with his own rational reflection about what sort of life he would like to lead and whether he has the powers necessary for it (e.g., whether it should be a trade, commerce, or a learned profession). (6:445)

The cultivation of one's own identity and way of life as a human being must, to a certain extent, be left up to the discretion of the individual in question, even as, in order to be virtuous, this telos must at the same time be guided by the maxim of moral self-perfection. And this point seems to be generalizable to moral self-perfection as such: "With regard to perfection as a moral end, it is true that in its end (objectively) there is only one virtue (as moral strength of one's maxims); but in fact (subjectively) there is a multitude of virtues, made up of several different qualities, and it would probably be impossible not to find in it some lack of virtue, if one wanted to look for it" (6:447). If 'it would probably be impossible not to find' in the multitude of virtues 'some lack of virtue', this indicates that there is no definitive answer to the question of how self-perfection can be universally attained. It must remain discretionary, and therefore up to the judgment of the individual moral agent—which, as I show in the next chapter, is the province not simply of applying rules or subsuming cases, but of one's independent formulation of particular maxims.

1.5   Personal autonomy, moral autonomy, and self-critique

This chapter has argued that the application presupposition, on which the moral principle of autonomy is to be transported from GMS and applied to MS, inadequately accounts for the developments in Kant's views post-1790—particularly the influence of the discovery of reflective judgment, with its own independent principle of autonomy. Thus, while recent literature has been right to draw attention to the changing status of autonomy post-GMS (Kleingeld 2017, forth.), the conclusion that the principle of autonomy drops out of Kant's mature thought altogether fails to account for the role it plays in Kant's final work, the OP. As I have demonstrated, the notion of autonomy remains central to the OP and consequently, given Kant's vision for the

crucial role this work plays in his philosophical system, to his considered position. However, it resurfaces in a new form: no longer defined in terms of self-legislation, but now a mediating concept that, through its activity of self-making, effectively generates the Kantian critical apparatus. One way to read Kant's insistence in MS on the necessity for individual judgment in applying practical principles to experience is that it implicitly presupposes a form of autonomy as self-making. As I have shown, Kant explicitly situates the role of practical, casuistical judgment as a specifically moral transition that would complement the transition represented by OP, and vice versa. Through Kant's elaboration of the imperfect duties of self-perfection, which can only be conceived teleologically, practical judgment emphasizes the necessity of reflective judgment for moral deliberation in general. Moreover, I have demonstrated that the comparison to natural teleology, which was already invoked in formulating FRE in GMS, introduces new nuance to Kant's initial wholesale rejection of perfectionism following his discovery of a new form of judgment for thinking the teleology of nature in KU.

The reflective dimension of practical judgment in these cases is subject to its own principle of autonomy, heautonomy, which governs the self-shaping of judgment. Indeed, I have attributed the fact that Kant defends a broader account of autonomy in OP than he does in the critical moral texts to Kant's initial articulation of the principle of heautonomy in KU. Thus, autonomy, on Kant's mature conception, incorporates *both* moral autonomy, or the practical dimension of moral deliberation that legislates a given maxim universally to all (and *ipso facto* to oneself), *and* heautonomy, or the reflective dimension, by which one produces new maxims whose scope extends only to oneself in practice.[45] In the next chapter, I elaborate on this distinction.

The consequences of this finding are twofold: first, autonomy remains indispensable to Kant's mature philosophical position. Second, the direction in which Kant develops his account of autonomy in this period puts into question the framing of the distinction between Kantian autonomy and contemporary accounts of autonomy referenced at the beginning of this chapter, or moral autonomy and personal autonomy respectively. If my interpretation has been convincing, Kant's account of autonomy as self-making is arguably more sophisticated than that of many contemporary theorists of (personal) autonomy, insofar as it provides us with a unified account of both morality and the creation and correction of principles of action (maxims) in general.

More centrally for my purposes in this dissertation, the reframing of this distinction suggests the possibility, currently underexplored, of *critiquing* our current

---

[45] On the distinction between universal legislation and self-given maxims, see Kleingeld and Willaschek (forth.) and Kleingeld (forth.).

constitution as subjects and practical agents. In order to answer to the demands of a contemporary account of autonomy, one which incorporates personal or political autonomy rather than limiting itself to moral autonomy alone, we need not take Korsgaard's route of appealing to 'practical identity' to *supplement* Kant's view, particularly given that this strategy prematurely forecloses important occasions for self-critique. After all, our identities are not wholly up to us and can thus often prove constraining, foreclosing productive possible avenues of action. Yet construing 'identity' so broadly as to encompass whatever it is a given agent happens to value, as Korsgaard (1996, 2009) does, precludes the possibility of *critique* of preexisting values. By contrast, Kant's account of the self-making of judgment is, on my interpretation, integrated into his account of practical maxim formation. Moreover, and as will prove key for my elaboration of the notion of moral conceivability in this dissertation, it integrates the possibility of critically evaluating contingent principles of action on the basis of experience. In the next chapter, I will address the role of judgment in maxim formation before turning to the issue of critique in later chapters.

Works cited

Allison, Henry E. 2001. *Kant's Theory of Taste: A Reading of the Critique of Aesthetic Judgment.* Cambridge, UK: Cambridge University Press.

Anscombe, G.E.M. 1956. *Mr. Truman's Degree.* Privately published pamphlet. Reprinted in *The Collected Philosophical Papers of G.E.M. Anscombe*, volume III: *Ethics, Religion and Politics*. Blackwell, Oxford: 1981, 62-71.

Bell, David. 1987. "The Art of Judgement." *Mind* 96: 382, 221-244.

di Giovanni, George. 1985. "Introduction: The Facts of Consciousness." In *Between Kant and Hegel: Texts in the Development of Post-Kantian Idealism*, edited by H. S. Harris and George di Giovanni, 2–50. State University of New York Press.

Flikshuh, Katrin. 2013. "Personal Autonomy and Public Authority." In Sensen (2013), 169-190.

Floyd, Juliet. 1998. "Heautonomy: Kant on Reflective Judgment and Systematicity." In *Kants Ästhetik*, edited by Herman Parret. Berlin: Walter de Gruyter. https://www.academia.edu/8669195/JF_1998_Heautonomy_in_Kant.

Förster, Eckart. 1989. "Kant's Selbstsetzungslehre." In *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*, edited by Eckart Förster. Stanford University Press.

———. 1990. "Fichte, Beck and Schelling in Kant's Opus Postumum." In *Kant and*

His Influence, edited by G. MacDonald Ross and Tony McWalter, 146–69. Thoemmes Continuum.

———. 1993. "Kant's Third Critique and the Opus Postumum." *Graduate Faculty Philosophy Journal* 16 (2): 345–58. https://doi.org/10.5840/gfpj199316221.

———. 2000. *Kant's Final Synthesis: An Essay on the Opus Postumum*. Harvard University Press.

———. 2012. *The Twenty-Five Years of Philosophy: A Systematic Reconstruction*. Translated by Brady Bowman. Cambridge, Mass: Harvard University Press.

Friedman, Michael. 1992. *Kant and the Exact Sciences*. Reissue edition. Cambridge, Mass.: Harvard University Press.

Ginsborg, Hannah. 2015. *The Normativity of Nature: Essays on Kant's Critique of Judgment.* Oxford, UK: Oxford University Press, 15-31.

Guyer, Paul. 1990. "Reason and Reflective Judgment: Kant on the Significance of Systematicity." *Noûs* 24 (1): 17–43. https://doi.org/10.2307/2215611.

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

———. 2013. "Progress toward Autonomy." In Sensen (2013), 71–86.

Herman, Barbara. 1993. *The Practice of Moral Judgment.* Cambridge, MA: Harvard University Press.

———. 2007. *Moral Literacy.* Cambridge, MA: Harvard University Press.

———. 2009. "Contingency in Obligation." *Nomos* 49, 17-53.

Hill, Thomas E. Jr. 1991. "The Importance of Autonomy." In *Autonomy and Self-Respect.* Cambridge: Cambridge University Press, 43-51.

Johnston, David. 1994. *The Idea of a Liberal Theory.* First Edition edition. Princeton, N.J: Princeton University Press.

Kleingeld, Pauline. 2017. "The Principle of Autonomy in Kant's Moral Theory: Its Rise and Fall." In *Kant on Persons and Agency.* Edited by Eric Watkins, 61–80. Cambridge University Press. DOI: 10.1017/9781316856529.005.

———. Forthcoming. "Freedom of the Will as Moral Self-Legislation." In *Proceedings of the 12th International Kant Congress*. Vienna.

Kleingeld, Pauline, and Marcus Willaschek. Forthcoming. "Kantian Autonomy without Self-Legislation of the Moral Law." In *Proceedings of the 12th International Kant Congress*. Vienna.

Klemme, Heiner F. 2013. "Moralized Nature, Naturalized Autonomy." In Sensen (2013): 193–211.

Korsgaard, Christine M. 1996. *The Sources of Normativity.* Cambridge: Cambridge

University Press.

———. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford /New York: Oxford University Press.

Longuenesse, Béatrice. 2001. *Kant and the Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Princeton, NJ: Princeton University Press.

Makkreel, Rudolf. 2002. "Reflective Judgment and the Problem of Assessing Virtue in Kant." *Journal of Value Inquiry* 36: 205-220.

Munzel, G. Felicitas. 1998. *Kant's Conception of Moral Character: The "Critical" Link of Morality, Anthropology, and Reflective Judgment.* Chicago, IL: Chicago University Press.

Neuhouser, Frederick. 1990. *Fichte's Theory of Subjectivity.* Cambridge: Cambridge University Press.

O'Neill, Onora. 1990. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139173773

———. 2016. "Kant on Indeterminacy, Judgment and Interpretation." De Gruyter Lecture. Chicago.

Rawls, John. 2000. *Lectures on the History of Moral Philosophy.* Edited by Barbara Herman. Cambridge: Harvard University Press.

Raz, Joseph. 1986. *The Morality of Freedom.* Oxford: Clarendon Press.

Recki, Birgit. 2001. *Ästhetik der Sitten: die Affinität von ästhetischem Gefühl und praktischer Vernunft bei Kant.* Vittorio Klostermann.

Sensen, Oliver (editor). 2013. *Kant on Moral Autonomy.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511792489.006.

Theunissen, Nandi. 2016 (online first 2013). "Kant's Commitment to Metaphysics of Morals." *European Journal of Philosophy* 24.1: 103-128. DOI: 10.1111/ejop.12051.

Thorndike, Oliver. 2018. *Kant's Transition Project and Late Philosophy: Connecting the Opus Postumum and Metaphysics of Morals.* Bloomsbury.

Wallner, Ingrid M. 1984. "A New Look at J. S. Beck's 'Doctrine of the Standpoint.'" *Kant-Studien* 75 (3): 294–316.

Wood, Allen. 1991. "Unsociable Sociability: The Anthropological Basis of Kant's Ethics." In *Philosophical Topics* 19.1: 325-251.

———. 1999. *Kant's Ethical Thought.* Cambridge, UK: Cambridge University Press.

———. 2003. "Kant and the Problem of Human Nature." In *Essays on Kant's Anthropology.* Edited by Brian Jacobs and Patrick Kain. Cambridge: Cambridge

University Press.

Wood, Allen and Robert Louden. 2012. "General Introduction." In *Lectures on Anthropology by Immanuel Kant*. Edited by Allen Wood, Robert Louden. Cambridge: Cambridge University Press, 1-10.

Zammito, John H. 1992. *The Genesis of Kant's* Critique of Judgment. Chicago: University of Chicago Press.