Kant on the Autonomy of Reason in the *Opus Postumum*
Sabina Vaccarino Bremner
University of Groningen

Introduction

Most Kant scholars have taken the Kantian principle of autonomy given in GMS (1785) to be indispensable to Kant's philosophy of the 1790s. Recent scholarship has challenged this presupposition, 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). In this paper, I argue that Kant *does* continue to invoke autonomy in the 1790s, albeit in OP rather than in the more explicitly moral texts (such as MS or RGV). Autonomy, as it is presented in OP, is no longer identified with the standpoint in GMS of the moral legislator freely conforming to the categorical imperative; instead, it is 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, who empirically realizes autonomy both by thinking and acting, and who constitutes herself as a rational agent in the process. In short, Kant comes to explicitly *broaden* his conception of autonomy, from moral autonomy to the autonomy of reason as such.

Scholars have attributed autonomy's shift in status to various factors: Kant's rejection of its governing metaphor of *a priori* legislation—or legislation without the empirical consent of the legislated (Kleingeld 2017), or Kant's new emphasis on moral autocracy rather than moral autonomy (Bacin, this volume).[1] Instead, I attribute the shift to a larger alteration in the conception of Kant's philosophical system; while this shift is one recognized by scholars of OP (Friedman 1992, Förster 2000)[2], its implications for autonomy, and Kant's moral philosophy more generally, have been neglected. In fact, Kant explicitly notes the influence of OP in the text of MS, in a passage in which he acknowledges the need for a parallel 'transition' in both texts

---

[1] While I find both claims plausible, I cannot go into deeper discussion of their merits here. Both suggest that Kant comes to emphasize the role of the *empirical* in moral reasoning, which is consonant with the broader shift I suggest here.

[2] Though see Emundts (2004) for a reading of OP as an extension of the critical project rather than a transformation of it.

between pure principles (whether metaphysical or practical) and empirical cases.[3] 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,[4] 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. So although Kant no longer conceives of autonomy as restricted to the moral domain, the new conception nevertheless remains relevant to Kant's mature moral philosophy.

This paper proceeds as follows: first, I present my reading of Kant's account of the autonomy of reason in OP*,* situating this account within Kant's general aims in this work. I then show that this conception constitutes a significant shift from Kant's better-known conception of autonomy as self-legislating the moral law. I conclude by presenting possible grounds for the shift: I propose that autonomy in OP incorporates Kant's conception of the autonomy (or self-directed normativity) *of judgment*—a notion christened 'heautonomy' in KU.

'Self-making' in OP

To contextualize my discussion, I will start by giving a brief summary of Kant's general aims in OP.[5] As scholars 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 MAN) 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" (Br 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

---

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

[4] See, e.g., Herman (2007), 133; Hill (1991), 43.

[5] 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 (1989, 1990, 1993), Friedman (1992), Emundts (2004).

other natural objects, and thinking beings who *create* the form of experience in which such laws are cognizable. If "experience… must be made" (OP 22:322, 108),[6] and "we have insight into nothing except what we can make ourselves" (22:353, 114), 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, 116). 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, 117). 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, 187).

It has been established that this position constitutes a departure from Kant's critical theoretical philosophy (Förster 1990; Wallner 1984; di Giovanni 1985). 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.[7] 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 (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.[8]

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,

---

[6] English translations of Kant's texts drawn from the Cambridge Edition of the Works of Immanuel Kant, with some modifications by the author. Due to Förster and Rosen's total restructuring of the Akademie Ausgabe rendering of OP in their Cambridge translation (see note 9), citations of OP include page number references to the Cambridge edition (Kant 1993).

[7] See also Kant's 1797 and 1798 letters to Tieftrunk (Br 12:207, 12:241), Garve (Br 12:255), and Kiesewetter (Br 12:258), in which Kant attacks these figures' popular new extensions of his critical project while (perhaps jealously) requesting news of how these new extensions are being received.

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

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, 189). 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, 189). 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, 189).

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, 212) 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, 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).

Autonomy and ideas 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). This concept of autonomy, 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:[9] "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,[10] particularly in describing the project of transcendental philosophy as such:

> Transcendental philosophy is the autonomy of ideas. (21:79, 246)

> 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)

> 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, 253)

In these passages, autonomy is associated with the production of *ideas*, and thus with *thinking* ideas rather than *acting* in accordance with them. 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

---

[9] One might situate this view not as a departure from Kant's critical position, but as a development of it. O'Neill (1989, 3-27) has written convincingly on the project of critique as a distinctly political or practical task. Moreover, in crucial passages of GMS and KrV, Kant presents practical reason, and therefore autonomy, as being evidenced by the cognitive subject's activity of self-consciousness (KrV A546/B574, GMS 4:457). However, as this paper aims to bring out, it is not until the post-critical period that this conception is more explicitly developed into autonomy as self-constitution (what we might now term 'personal autonomy'), rather than strictly *moral* autonomy. On the personal/moral autonomy distinction in contemporary political philosophy, see Raz (1986, 370), Johnston (1994, 75), and Flikshuh (2013).

[10] 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. However, the precise ordering doesn't much matter for my purposes.

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,[11] attributing a role to maxims in *theoretical* cognition in allowing cognizers to proceed in accordance with ideas that would otherwise could not be cognizable. 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 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.

Heautonomy and connection to moral reasoning

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, 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).[12]

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

---

[11] See, e.g., KrV A648/B676, A651-2/B680-1, A666/B694, A680/B708; KU 20:205, 20:219, 5:182, 5:184.
[12] See Floyd (1998, 205).

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. Thus, if Kant's characterization of reflective judgment in KU refers to cases that are *merely* [*bloß, nur*] reflective (rather than reflective *and* determining, or reflective *and* practical[13]), 'autonomy', as the more general concept, will also comprise 'heautonomy'. Moral autonomy presupposes heautonomy, while theoretical spontaneity becomes more explicitly normative, or 'heautonomous': both moral and theoretical reasoning require the subjective, self-given generation of maxims. 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*]" (OP 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, 249). 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)

---

[13] Many scholars have argued that theoretical account of cognition requires both reflective and determining judgment; see, e.g., Bell (1987), Floyd (1998), Guyer (1990), Allison (2001, 20), Ginsborg (2015). In particular, see Longuenesse (2001, 164) for analysis of Kant's references to '*merely* reflecting judgment' in KU. O'Neill (1989, 36), Munzel (1998), Recki (2001), Makkreel (2002), and Klemme (2013) each suggest that reflective judgment may also be incorporated into practical deliberation.

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. Kant characterizes the heautonomous principles of judgment in the Second Introduction of KU as "themselves fit neither for theoretical nor for practical use", thus "mak[ing] possible a transition" from one domain to the other (5:176).

Taking stock

This paper has shown that, while scholars are right to draw attention to the changing status of autonomy post-GMS*,* it is mistaken to conclude that the notion of autonomy drops out of Kant's mature thought altogether. As I have demonstrated, the notion of autonomy remains central to Kant's final work, 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. In this process, judgment is subject to its own principle of autonomy, heautonomy, which governs the self-shaping of judgment. Finally, I have proposed one way to make sense of why autonomy is conceived as the autonomy *of reason* in OP by tracing this conception back to Kant's initial articulation of the principle of heautonomy in KU. On my proposal, which requires future development, 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.[14]

---

[14] On the distinction between universal legislation and self-given maxims, see Kleingeld (2018) and Kleingeld and Willaschek (2019).

Works cited

Allison, Henry (2001): *Kant's Theory of Taste: A Reading of the Critique of Aesthetic Judgment*. Cambridge: Cambridge.

Bell, David (1987): "The Art of Judgement". In: *Mind* 96.382, 221-244.

di Giovanni, George (1985): "Introduction: The Facts of Consciousness". In: H.S. Harris/George di Giovanni, (Eds.): *Between Kant and Hegel: Texts in the Development of Post-Kantian Idealism.* Albany: SUNY, 2–50.

Emundts, Dina (2004): *Kants Übergangskonzeption im Opus Postumum.* Berlin: de Gruyter.

Flikshuh, Katrin (2013): "Personal Autonomy and Public Authority". In: Sensen, Oliver (Ed.): *Kant on Moral Autonomy*. Cambridge: Cambridge, 169-190.

Floyd, Juliet (1998): "Heautonomy: Kant on Reflective Judgment and Systematicity". In: Parret, Herman (Ed.): *Kants Ästhetik*. Berlin: de Gruyter, 192-218.

Förster, Eckart (1989): "Kant's Selbstsetzungslehre". In: Förster, Eckart (Ed.): *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum.* Palo Alto: Stanford.

Förster, Eckart (1990): "Fichte, Beck and Schelling in Kant's Opus Postumum". In: G. MacDonald Ross/Tony McWalter (Eds.): *Kant and His Influence*. Bristol: Thoemmes Continuum, 146-169.

Förster, Eckart (1993): "Kant's Third Critique and the Opus Postumum". In: *Graduate Faculty Philosophy Journal* 16.2, 345–358.

Förster, Eckart (2000): *Kant's Final Synthesis: An Essay on the Opus Postumum*. Cambridge: Harvard.

Friedman, Michael (1992): *Kant and the Exact Sciences*. Cambridge: Harvard.

Ginsborg, Hannah (2015): *The Normativity of Nature: Essays on Kant's Critique of Judgment.* Oxford: Oxford.

Guyer, Paul (1990): "Reason and Reflective Judgment: Kant on the Significance of Systematicity." In: *Noûs* 24.1: 17–43.

Herman, Barbara (2007): *Moral Literacy*. Cambridge: Harvard.

Hill, Thomas E. Jr (1991): *Autonomy and Self-Respect*. Cambridge: Cambridge.

Johnston, David (1994): *The Idea of a Liberal Theory*. Princeton: Princeton.

Kant, Immanuel (1993): *Opus Postumum.* Eckart Förster, M. Rosen (Eds), Eckart Förster (trans.). Cambridge: Cambridge.

Kleingeld, Pauline (2017): "The Principle of Autonomy in Kant's Moral Theory: Its Rise and Fall." In: Watkins, Eric (Ed.): *Kant on Persons and Agency*. Cambridge: Cambridge.

Kleingeld, Pauline (2018): "Moral Autonomy as Political Analogy: Self-Legislation in Kant's Groundwork and the Feyerabend Lectures on Natural Law (1784)." In: Stefano Bacin/Oliver Sensen (Eds.): *The Emergence of Autonomy in Kant's Moral Philosophy*. Cambridge: Cambridge,158-175.

Kleingeld, Pauline, and Marcus Willaschek (2019): "Autonomy Without Paradox." In: *Philosopher's Imprint* 19.6, 1-18.

Klemme, Heiner F (2013): "Moralized Nature, Naturalized Autonomy." In: Sensen, Oliver (Ed.): *Kant on Moral Autonomy*. Cambridge: Cambridge, 193–211.

Korsgaard, Christine (2009): *Self-Constitution: Agency, Identity, and Integrity*. Oxford, New York: Oxford.

Longuenesse, Béatrice (2001): *Kant and the Capacity to Judge*. Princeton: Princeton.

Makkreel, Rudolf (2002): "Reflective Judgment and the Problem of Assessing Virtue in Kant." In: *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: University of Chicago Press.

O'Neill, Onora (1989): *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge.

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

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

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

Wallner, Ingrid (1984): "A New Look at J. S. Beck's 'Doctrine of the Standpoint.'" In: *Kant-Studien* 75.3, 294–316.