



The Principle of Autonomy's Enduring Validity

Marie Newhouse¹ 

Received: 29 January 2023 / Accepted: 1 March 2023 /
Published online: 10 April 2023
© The Author(s) 2023

Abstract

Pauline Kleingeld has argued persuasively that Kant's Principle of Autonomy draws an analogy between two relationships: 1) that between an individual agent and their maxim, and 2) that between a legislator and their legislation. She also suggests that Kant's evolving views on the normative significance of popular elections made his analogy inapt, which explains its disappearance from his later writings. This comment concurs with Sorin Baiasu that the merits of Kant's analogy were untouched by his evolving political views. The analogy presupposes that when lawmakers legislate, they do so as representatives of the people, but even Kant's mature writings posit that unelected lawmakers are representatives of the people in the necessary sense. Facts about whether a lawmaker acquired legislative authority by election, heredity, or conquest do have normative implications in the larger context of Kant's political theory, but such facts are necessarily temporally upstream of the relation between legislator and legislation, just as facts about how an agent became rational are upstream of the relation between individual agent and maxim. Because the analogy focuses exclusively on these two relationships, it cannot be damaged by such upstream facts. Moreover, we should consider the possibility that Kant's Principle of Autonomy compares a human will to the united general will of the people itself rather than to the will of a fallible public official or institution. On this alternative reading, it is even more clear that Kant's evolving views on the significance of elections could have no possible relevance to the Principle of Autonomy.

Keywords Kant · Pauline Kleingeld · Principle of Autonomy · Categorical Imperative · Moral philosophy · Analogy

✉ Marie Newhouse
m.newhouse@surrey.ac.uk

¹ School of Law, University of Surrey, Guildford, Surrey GU2 7XH, UK

1 Introduction

Kleingeld has offered a promising interpretation of the *Groundwork's* Principle of Autonomy, according to which Kant is analogizing (1) an individual agent's relation to her maxim to (2) a legislator's relation to legislation. (Kleingeld, 2018a) Baiasu praises this well-supported interpretation, and he finds Kleingeld's distinction between the Principle of Autonomy and the will's property of autonomy illuminating, as do I. (Baiasu, 2023) However, he expresses skepticism about her further argument that Kant's evolving views on the significance of elections undermined the analogy encapsulated in the Principle of Autonomy and can therefore explain the latter's absence in the *Metaphysics of Morals*. (Kleingeld, 2018b) In what follows, I explain why I find Baiasu's argument persuasive and offer additional considerations in support of his thesis. Then, I suggest that popular English translations of Kant's Formula of Autonomy have obscured a possibility that we should consider: Kant may have intended to compare a human will to the united general will of the people itself rather than to the will of a public official or institution. If this alternative reading were correct, then Kant's evolving views on the significance of elections could have no possible relevance to the Formula of Autonomy.

2 Considerations in Support of Baiasu's Thesis

Kleingeld and Baiasu disagree about whether a new normative requirement that legislators be popularly elected would have ruined the analogy captured by the Principle of Autonomy. To evaluate their arguments, I will clarify what I take to be Kleingeld's account of this analogy. Kleingeld insightfully observes that Kant's original "modal content" criterion for the justice of laws – possible universal adoption – is "formally the same" as the criterion set out in the Categorical Imperative's Formula of Universal law. (Kleingeld, 2018a: 167) The relation between agent and maxim is thus identical to the relation between legislator and legislation: both pairings are interpolated by the same formal principle, which in both cases identifies agents' subjective principles (maxims or legislation) as morally permissible or prohibited. See Fig. 1.

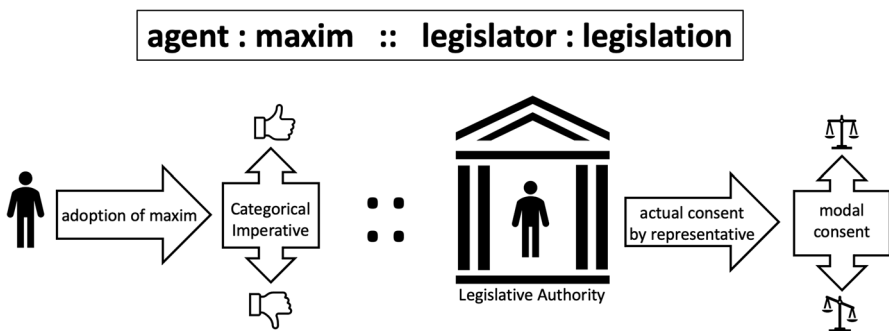


Fig. 1 Kleingeld's account of the principle of autonomy as political analogy

Kleingeld suggests that this analogy is absent in Kant's later work because he subsequently adopted an additional criterion for the justice of legislation: "actual consent by the citizens, via their elected representatives". (Kleingeld, 2018b: 73) Her language seems to identify two separate requirements: (1) citizens' actual consent to legislation via their representative(s); and (2) popular election of those representatives.

Baiasu argues that (1) is not a change to Kant's political theory. (Baiasu, 2023) I find this position persuasive on textual grounds. An unelected king literally consents to legislation in the act of issuing it, and Kant believed that such a king represents the people when he does so. For example, the Feyerabend lectures refer to "the individual person" who wields state sovereignty as "*repraesentant* of the people." (FL 27:1382) ¹ In other passages dated between 1785 and 1789, Kant refers to a monarch as "one to whom [the people] have merely transferred the right to act as their representative" and as "only the representative of the people". (RPR 19:593) In a later remark, Kant refers to Louis XVI as the "king [who] otherwise represented the people" before he abandoned his sovereignty to be claimed by the French National Assembly. (RPR 19:595–96) Actual consent by the people via their representative(s) was therefore always a feature of Kant's legislative process.

In reply to Baiasu, Kleingeld might point out that consent by the people (*des Volks*) is not necessarily the same thing as consent by citizens (*Staatsbürger*). This would be a fair point, and Kant muddies the water by using the same word to refer both to "active citizens" (who have the right to vote) and to "passive citizens" (who do not). (MM 6:314) But in what Kleingeld identifies as the key passage, Kant refers to "all the citizens united", which arguably suggests that he means to refer to passive and active citizens collectively—a group coextensive with "the people". (MM 6:341) This reading is plausible because Kant uses the term "citizen" in this inclusive sense elsewhere. For example: "no human being in a state can be without any dignity, since he at least has the dignity of a citizen." (MM 6:329) Moreover, it is hard to see why Kant would have singled out active citizens in this passage given that he is discussing legislating *through representatives* rather than directly. Since both active and passive citizens are represented by lawmakers, lawmakers who literally consent to legislation would be doing so on behalf of both groups.

¹ In what follows, in citing Kant's works, the following abbreviations are used:

AA: German edition of Kant's complete works (*Kants gesammelte Schriften*) (1900-).

G: *Groundwork of the metaphysics of morals* (*Grundlegung zur Metaphysik der Sitten – AA 04*) (1785), in Kant (1996a, b, c, d:43–108).

MM: *The metaphysics of morals* (*Die Metaphysik der Sitten – AA 06*), comprising the *Metaphysical First Principles of the Doctrine of Right* (*Metaphysische Anfangsgründe der Rechtslehre*) (1797) and the *Metaphysical First Principles of the Doctrine of Virtue* (*Metaphysische Anfangsgründe der Tugendlehre*) (1797), in Kant (1996a, b, c, d: 353–603).

PP: *Toward perpetual peace* (*Zum ewigen Frieden*) (1795), in Kant (1996a, b, c, d: 317–51).

RPR: *Reflections on the Philosophy of Right* (*Reflexionen zur Rechtsphilosophie – AA 19*), in Kant (2016a, b: 7–72).

FL: *Natural Right lecture notes Feyerabend* (*Naturrecht Feyerabend* (Winter 1784) – AA 27), in Kant (2016a, b: 81–180).

I am using the translations listed in the Bibliography.

Baiasu then argues that although (2) would be a change to Kant's political theory, it would not alter the relation between legislator and legislation, which is the focus of Kant's analogy. (Baiasu, 2023) His conclusion on this point seems logically sound. Whether elected or not, an agent must hold the legislative authority in order to legislate at all, just as an agent must be rational in order to freely adopt a maxim, and Kleingeld accepts that Kant at all times considered unelected leaders capable of holding the legislative authority. (MM 6:319) Facts about whether a lawmaker acquired that authority by election, heredity, or conquest may have normative implications in the larger context of Kant's political theory, but such facts are necessarily temporally upstream of the relation between legislator and legislation, just as facts about how an individual became rational are upstream of the relation between an individual and her maxim. Modal consent therefore remains the only normative principle in the *Metaphysics of Morals* that interpolates legislator and legislation, a fact that preserves Kant's analogy unscathed despite his evolving views about the significance of elections.

Baiasu also considers and rejects an indirect potential threat to the soundness of Kant's analogy: the inference that unelected lawmakers give laws monologically (that is, without interpersonal deliberation), whereas elected lawmakers have a dialogical lawmaking process. Baiasu denies that this change would ruin Kant's analogy because the *Groundwork*-era Kant describes individual decision-making in dialogic terms as "a deliberative process in which the will or other (alien) impulses may give laws", which suggests functional equivalence. (Baiasu, 2023) I find this argument especially plausible because Kant elsewhere conversely likens a political election (a dialogic process) to individual practical deliberation: "In voting society is considered to be one person and the members to be particular wills and inclinations of that one person." (RPR 19:447) I find Baiasu's conclusion on this point sound for a separate reason as well: election of lawmakers is neither a necessary nor a sufficient condition for a dialogic lawmaking process. Kant's "aristocratic" form of government, in which a hereditary group agent may exercise the legislative authority dialogically, demonstrates that it is unnecessary. (PP 8:294f; MM 6:339) Kant's claim that an individual monarch can become sovereign "through election" suggests that it is not sufficient. (FL 27:1388) Therefore, we cannot assume that a shift from unelected to elected lawmakers entails a shift from a monologic to a dialogic process in the first place.

3 Tweaking Kleingeld's Interpretation of Kant's Formula of Autonomy

So far, I have assumed that Kleingeld is correct to think that Kant intends to compare a private human will to the will of a public official or institution such as the UK Parliament. But this is not necessarily correct. Kant might have meant to draw an analogy between (1) a private human will and its maxim and (2) the united general will of the people and its law. In that case, Kant's evolving views about the significance of elections could not possibly threaten his analogy because elections only concern the selection of empirical lawmakers.

An examination of Kant’s original German text lends plausibility to this alternative reading. Kant articulates his Formula of Autonomy twice:

- 1) der Idee des Willens eines jeden vernünftigen Wesens als *allgemein-gesetzgebenden Willens*
- 2) das *Prinzip* eines jeden menschlichen Willens, als *eines durch alle seine Maximen allgemein gesetzgebenden Willens* (G 4:432)

The word *allgemein* can be translated as “general” or “universal”. The word *gesetzgebenden* can be translated as “legislative” or “lawgiving”. In both German formulations, *allgemein gesetzgebenden* is an adjective phrase modifying the noun *Willens* (will). However, both Mary Gregor and Kleingeld convert these words into a verb clause. Gregor translates the two phrases above as (1) “the idea of the will of every rational being as a will giving universal law” and (2) “the principle of every human will as a will giving universal law through all its maxims”. (G 4:432) Kleingeld concurs with Gregor on Kant’s first statement but offers a different translation of his second: the “principle of every human will as a will that is universally legislating through all its maxims”. (Kleingeld, 2018a: 160)

In each of the above English translations, the decision to recast *allgemein gesetzgebenden* as a verb clause creates a linguistic asymmetry between the two halves of Kant’s analogy. The first half describes an individual will in terms of its quality (“rational” or “human”) and does not refer to its activity. The second half describes the activity of the analogous will (“giving universal law” or “legislating universally”) but does not refer to its quality. This linguistic mismatch alone raises a doubt about whether these translations convey Kant’s precise idea. Moreover, this translational choice sharpens a substantive ambiguity: a will engaged in the activity of giving universal law (or legislating universally) could be that of a public official or institution, but it could also be the united general will itself.

A more literal translation of Kant’s original German text—one that treats the words *allgemein gesetzgebenden* as an adjective phrase describing the will to which an individual human will should be likened—suggests the latter possibility. Adopting this approach, Kant’s two statements of the Formula of Autonomy could be translated like this:

- 1) the idea of the will of every rational being as a universally-legislative will.
- 2) the principle of every human will as a (through all its maxims) universally legislative will.

These English renderings are awkward, but they intriguingly suggest that the second half of Kant’s analogy might refer to the will of a constructive rather than empirical agent. Similar language in the contemporaneous Feyerabend lectures lends some plausibility to this reading. In them, Kant is recorded referring to *allgemeinen Willens des Volks*, which Frederick Rauscher and Kenneth R. Westphal translate as “general will of the people”. (FL 27:1383) The lectures also include a reference to *dem allgemeinen Willen*, which Rauscher and Westphal render as “the

universal will”. In both contexts, these words seem to refer to a constructive rather than empirical agent. (FL 27:1384)

It might initially seem as though my proposed tweak to Kleingeld’s interpretation replaces a linguistic asymmetry with a more serious philosophical asymmetry. As Fig. 1 shows, both private moral agents and public officials can adopt principles of action (either maxims or legislation) that fail to conform to the Categorical Imperative or the modal consent test, respectively. By contrast, the united general will itself “can do absolutely no wrong” and is therefore not relevantly identical to a human agent. (FL 27:1382) But this apparent worry dissipates when one recalls that Kant’s Formula of Autonomy is supposed to be an *imperative*. On the reading I am proposing, Kant would not be claiming that both private moral agents and public officials can in fact either succeed or fail to meet applicable norms. Rather, he would be exhorting human agents to emulate the united general will by adopting only maxims that can serve as universal laws.

4 Conclusion

Kleingeld deserves praise for her well-supported and thought-provoking analysis of the Formula of Autonomy. Her proposed account of Kant’s analogy seems to me to be at least *very* close to correct, and the reading I suggest above would constitute a minor clarification of her interpretation. Paired with her rigorous supporting analysis of the concept of self-legislation, it contributes significantly to our understanding of the *Groundwork*. If Kant’s analogy refers to empirical lawmakers, then Baiasu’s arguments for its continuing validity appear sound for the reasons I have offered above. If my speculation that the analogy in fact refers to the united general will itself were correct, then its enduring validity would be even more unassailable.

Acknowledgements I am grateful to Ralf M. Bader, Sorin Baiasu, Christopher P. Taggart, Jens Timmermann, and Allen W. Wood for valuable advice and feedback. All errors are my own.

Declarations

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baiasu, S. (2023). How Far Do We Self-Legislate? In *Philosophia*.
- Kant, I. (1996a). Groundwork of The metaphysics of morals. In *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy* Ed. and tr. M. J. Gregor. Cambridge University Press.
- Kant, I. (1996b). On the common saying: That may be correct in theory, but it is of no use in practice. *The Cambridge Edition of the works of Immanuel Kant: Practical philosophy* (Ed and tr. M. J. Gregor). Cambridge University Press.
- Kant, I. (1996c). Toward perpetual peace. In *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy* (Ed. and tr. M. J. Gregor). Cambridge University Press.
- Kant, I. (1996d). The metaphysics of morals. In *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy* (Ed. and tr. M. J. Gregor). Cambridge University Press.
- Kant, I. (2016a). Reflections on the philosophy of right. In *The Cambridge Edition of the works of Immanuel Kant: Lectures and drafts on political philosophy* (Ed. F. Rauscher and tr. F. Rauscher, & K. R. Westphal). Cambridge University Press.
- Kant, I. (2016b). Natural right course lecture notes by Feyerabend. In *The Cambridge Edition of the works of Immanuel Kant: Lectures and drafts on political philosophy* (Ed. F. Rauscher and and tr. F. Rauscher, & K. R. Westphal). Cambridge University Press.
- Kleingeld, P. (2018a). Moral Autonomy as Political Analogy: Self-legislation. In Kant's Groundwork and the Feyerabend Lectures on Natural Law (1784). In S. Bacin & O. Sensen (Eds.), *In The Emergence of Autonomy in Kant's Moral Philosophy* (pp. 158–75). Cambridge University Press.
- Kleingeld, P. (2018b). The Principle of Autonomy in Kant's Moral Theory: Its Rise and Fall. In E. Watkins (Ed.), *In Kant on Persons and Agency* (pp. 61–79). Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.