

“Under the Guise of the Good”: Kant and a Tenet of Moral Rationalism

Stefano Bacin (Milano)

in Violetta Waibel, Margit Ruffing, David Wagner (eds.), *Natur und Freiheit. Akten des 12. Internationalen Kant-Kongresses*, Berlin–Boston, De Gruyter, 2018, 1705-1714

— please cite from published version —

According to an ancient thought in philosophy of action and moral psychology, every instance of rational volition happens “under the guise of the good”, as it has become customary to translate the formula *sub ratione boni* introduced by Aquinas. The thought goes back to Aristotle and through the whole history of moral philosophy, and is still an important topic of discussion. David Velleman and Kieran Setiya recently put forward strong criticisms of that view, while Joseph Raz, Sergio Tenenbaum and others have defended it. Both in the historical debates and in recent discussions, the Guise of the Good Thesis (hereafter: the Thesis) represents a genuine dogma of rationalism in moral philosophy.¹ Many influential commentators have maintained that Kant belongs in that champ too, even that he “*explicitly* endorses” the Thesis.² Those commentators also regard the Thesis as playing a crucial role in important arguments in Kant’s moral works, such as

¹ See the extended critique in Setiya, Kieran: *Reasons without Rationalism*. Princeton 2007.

² Herman, Barbara: *The Practice of Moral Judgement*. Cambridge 1993, 208f. fn. Cf. e.g. Korsgaard, Christine M.: *Creating the Kingdom of Ends*. Cambridge–New York 1996, 115f.; Engstrom, Stephen: *The Form of Practical Knowledge*. Cambridge 2009, 124ff. Cf. also Forman, David: *Appetimus sub ratione boni: Kant’s Practical Principles between Crusius and Leibniz*. In: *Kant und die Philosophie in weltbürgerlicher Absicht*. Ed. Stefano Bacin, Alfredo Ferrarin, Claudio La Rocca, Margit Ruffing. Berlin–Boston 2013, vol. 5, 323-334.

in *GMS* III or in the justification of the Formula of the End in Itself.³ While this understanding of Kant's conception has hardly been opposed so far,⁴ only recently some interpreter has felt the need to defend the attribution of the Thesis to Kant in some detail.⁵

In addition to the philosophical appeal of Thesis,⁶ attributing it to Kant is appealing for specific reasons, the most important one being that it offers an elegant answer to the question of the foundation of moral obligation. On this view, "commitment to moral principle is built into rational choice *per se*,"⁷ so every rational volition would entail a (formal) tendency towards the moral law. Still, the welcome philosophical outcome does not make the interpretation right as a reading of Kant's view. In this respect I find it not entirely convincing. Not only the textual evidence does not appear conclusive to me (as is also granted by some advocate of this view),⁸ since only quite early passage directly supports the Thesis (AA 01:400f.), but the interpretation attributing the Thesis to Kant amounts to a dubious understanding of the relationship of Kant's moral philosophy to previous views. Kant would then maintain the same fundamental view of rational action as prior moral rationalists, even within a different account of the mind and morality. Here I take this historical angle to very briefly explore under which conditions Kant can be taken to be that close to previous moral rationalism on this point. He might simply have held a different version of the Thesis, but then also the differences, if there are some, should be significant.

1

³ On the role of the Thesis in the derivation of the Formula of the End in Itself see the critical remarks in Kerstein, Samuel: *Deriving the Formula of Humanity*. In: *Groundwork for the Metaphysics of Morals*. Ed. Christoph Horn, Dieter Schönecker. Berlin-New York 2006, 200-221.

⁴ See chiefly Hill, Thomas E., Jr.: "Personal Values and Setting Oneself Ends," in Hill, *Human Welfare and Moral Worth*. Oxford-New York 2002.

⁵ See Reath, Andrews: *Did Kant Hold that Rational Volition is Sub Ratione Boni?* In: *Reason, Value, and Respect*. Ed. Mark Timmons, Robert N. Johnson. Oxford-New York 2015, 232-255. See also Reath's paper published in the present proceedings.

⁶ See Orsi, Francesco: *The Guise of the Good*. In: *Philosophy Compass* 10 (2015), 714-724.

⁷ Reath, *Did Kant Hold*, 232.

⁸ Cf. Reath, *Did Kant Hold*, 235.

After Aristotle and Aquinas, other writers, historically much closer to Kant, regarded the Thesis as a central piece of their account of rationality. Comparing Kant's view with theirs might be more helpful to shed light on the issue.⁹ If Leibniz defended what Locke had called the “traditional” view, namely, that what we consider “the greater good is that alone which determines the will”,¹⁰ Wolff insisted on the Thesis even more, presenting it as the fundamental “law of desire (*lex appetitus*)”, which he states again and again.¹¹ In Wolff’s words, “we cannot will anything that we do not hold to be good, and cannot not will without what we view as evil”.¹² Note that, as Wolff remarks, ‘law’ means here a principle describing a *factual* regularity, like a physical law. Defending Wolff against Lange, a Wolffian even calls the principle “a rule held by every rational philosophy and confirmed by everyday experience”.¹³

When Kant mentions the “old formula of the Schools” (KpV AA 05: 59), he refers to this background. Mentioning it is not the same as endorsing it, though. Critical distance and the attempt at integrating a partial insight in a new perspective merge in Kant's remark. Kant’s attitude towards the “old Scholastic principle” discussed in § 12 of *KrV*, *quodlibet ens est unum, verum, bonum*, is analogous to his view on Wolff’s “law of desire”. He rejects both as long as they is understood as “the School” does, but proposes a corrected version that should maintain the grain of truth they contain.

The principle *quodlibet ens etc.* is relevant here also because pre-Kantian authors holding the Thesis refer to *bonum* as a transcendental notion. In this perspective, everything is ‘good’, and all that morally matters is to discriminate between real (or full-fledged) goodness and apparent (i.e. insufficient) goodness. Wolff’s perfectionism is a clear example of this paradigm. The traditional view has, thus, a metaphysical background

⁹ Only Forman, *Appetimus*, has taken the views of Kant's predecessors into consideration so far.

¹⁰ Locke, John: *An Essay Concerning Human Understanding* (1690). Ed. P. H. Nidditch. Oxford 1975, IX.21.29.

¹¹ Wolff, Christian: *Psychologia empirica* (1738). Hildesheim 1968, § 904. Cf. e.g. *Oratio de Sinarum philosophia practica* (1726). Ed. M. Albrecht. Hamburg 1985, 32f..

¹² Wolff Christian: *Vernünfftige Gedancken von Gott, der Welt und der Seele des Menschen [...]* (1751⁴). Hildesheim 1983), § 506. Cf. *Anmerkungen*, § 162.

¹³ Carpov, Jakob: *Ausführliche Erläuterung der Wolffischen vernünfftigen Gedancken von der Menschen Thun und Lassen*. Franckfurt–Leipzig 1735, 75, § 49.

that Kant cannot accept. Wolff's thought expresses both a metaphysical and an epistemological continuity: For Wolff, namely, what differentiates rational desire (or aversion) from sensible desire is merely that the former arises not from confused, but from distinct representations. Apart from that, *there is no difference in the content*, as both sensible and rational desire "tend to the same object".¹⁴

Kant, thus, cannot "aims to be in conformity with the "old formula of the schools" and thus with Wolff's *lex appetitus*",¹⁵ which has a metaphysical backdrop incompatible with his general philosophical view. In fact, Kant's treatment of the old formula displays voluntarist elements that give to his view a different turn. Kant's remarks on the formula are closer to Crusius' criticism of the Wolffian version of the Thesis than to Wolff's view. Crusius' view that the good consists in a relation to the will entails that the will cannot be understood as a capacity to will something previously cognised as good.¹⁶ Two relations must be distinguished: what Crusius calls 'metaphysically good' is by definition determined by the human will, while the morally good, determined by the divine will, which is not actually willed, but is what the human will *ought to* will.¹⁷ Kant's remarks on the "old formula" are closer to this view than to Wolff's, as Kant both maintains that 'good' means a relation to the will and distinguishes kinds of good, depending on whether that relation holds a priori or a posteriori (cf. KpV AA 05: 60) and has descriptive or prescriptive nature. A first difference, thus, separating Kant from the traditional Thesis is that in Kant's view there is no continuity in the *object* of the will.

2

Kant, however, goes beyond Wolff's and Crusius' views, since, if he does not hold that a volition follows from a more or less clear representation of something as good, he does not maintain that the volition causes its object to be good either. In Kant's view, in fact, volition is governed not by knowledge or belief, but by principles. Indeed, one might construe Kant's view as a version of the Thesis in terms of principles. This comes closer to

¹⁴ Wolff, *Psychologia empirica*, § 908.

¹⁵ Forman, *Appetitus*, 331.

¹⁶ Crusius Christian August: *Anweisung vernünftig zu leben* (1744). Hildesheim 1969, § 26.

¹⁷ Crusius, *Anweisung*, § 27.

what Andrews Reath has recently suggested. On his reconstruction, “rational volition is based on *practical reasoning* aimed at judgments of goodness that make a tacit claim to universality”.¹⁸ Yet, the turn from representations of beliefs to principles marks a significant change by itself. In the traditional view, a belief-desire conception of volition is at the centre, that is, volition is thought of as a product of a belief about a property of an object that generates a desire. This model holds both for Wolff and for Raz (to put it roughly), but not for Kant, who rather holds that volition draws on principles that determine possible actions as actions that one should do. Two conceptions of practical cognition diverge, here. According to Wolff’s recognitional view, practical cognition is about objects and their properties, i.e. about *something* regarding which we have beliefs that can be more or less certain and accurate. On the contrary, for Kant practical cognition is rather about actions. Strictly speaking, goodness cannot be regarded as a property of objects or state of affairs, but only applies to actions as resulting from maxims (cf. KpV, AA 05: 60), that is, goodness is *not* analogous to a natural property.

3

So far I have highlighted fundamental differences between the traditional version of the Thesis and Kant’s view. However, recent interpretations might intend to ascribe to Kant a version of the Thesis which does not commit him to any of the assumptions I have pointed at. Wolff does hold that “every rational volition entails a commitment to moral values”, to borrow Reath’s phrase, since in Wolff’s view there is a *material* continuity between different degrees of goodness as degrees of the same scale. Analogously, Kant’s view can be interpreted as centred on a *formal* continuity through the different forms of rational volition. Reath hence claims that “what makes an act the kind of rational activity it is [...] that it understands itself to have a certain formal aim”.¹⁹

Yet, such a reading runs against Kant’s rejection of a continuity between sensible and intellectual faculties. While the Thesis emphasises the structural homogeneity of morally adequate and morally defective choice, Kant’s view stresses the *structural* difference

¹⁸ Reath, *Did Kant Hold*, 234.

¹⁹ Reath, *Did Kant Hold*, 249.

between two ways of determining the will. Because there is no continuity between empirically conditioned and pure determinations of the will, Kant does not point out any homology between them and speaks, instead, of a “contest” opposing higher and lower faculty of desire (“*lucta facultatis appetitivae superioris atque inferioris*”: AA 28: 678; cf. AA 28: 256).²⁰

The basic thought in ascribing to Kant a formal version of the Thesis, however, lies in the underlying idea that every determination of the will as such entails a claim to universality. In Reath’s formulation, “rational volition is based on practical reasoning aimed at *judgments of goodness* that make a *tacit claim to universality*”.²¹ This interpretation encounters some difficulties, though. In Kant’s view, the defining feature of rational agency is that it draws on maxims, as opposed to stimuli (cf. e.g. AA 28: 678), but Kantian maxims are neither “judgments of goodness” nor “make a tacit claim to universality”. I shall now briefly comment on these two aspects in turn.

First, Kantian maxims are not propositions or rules about the good, unlike Wolffian maxims.²² In fact, whereas Wolff understood maxims as general judgments about the value properties that we have cognised in some object, Kant never defines maxims through this feature and merely conceives of them as subjective principles of willing (cf. GMS, AA 04: 400 fn.), which are not regarded as guided by an antecedent representation or belief about morally relevant properties of any object. Maxims are “practical *rule[s]* that reason determines in conformity with the conditions *of the subject* (quite often his ignorance, or his inclinations), and [are] thus the principle[s] according to which the subject *acts*” (GMS, AA 04: 420f. fn.). That is, maxims entail no normative claim about how the subject *ought to act* (or should act) to correspond to the (moral or instrumental) goodness envisaged through an evaluative representation of its object. Kantian maxims are not as such principles of evaluation, as the Thesis would require. Moreover, if any maxim is supposed to amount to a statement to the effect that the agent holds something

²⁰ See on that Wuerth, Julian: *Kant on Mind, Action, and Ethics*. Oxford 2014, 230f.

²¹ Reath, *Did Kant Hold*, 234.

²² Cf. Wolff, Christian: *Vernünfftige Gedancken von der Menschen Thun und Lassen [...]* (1733). Ed. H. W. Arndt. Hildesheim–Zürich–New York 1996, § 190-192; *Philosophia practica universalis [...], Pars posterior [...]* (1739). Hildesheim–New York 1979, § 829f.

to be good, then the idea of a maxim of maxims introduced to point at the origin of the evil appears particularly difficult to grasp (cf. RGV, AA 06: 23f. and 31.)

Second, not every maxim as such aims at universality, but it is merely a principle stating a determination of the will regarded as subjectively valid (cf. KpV AA 05: 19). Maxims can thus very well include exceptions to principles, without aiming at universal validity, as one of Kant's examples of maxims explicitly shows: "when I believe myself to be in need of money I shall borrow money, and promise to repay it, even though I know that it will never happen" (GMS, AA 04: 422). In fact, "if we now attend to ourselves in every transgression of a duty, we find" not merely that we cannot will, but "that *we actually do not will* that our maxim should become a universal law" (GMS, AA 04: 424).

In Reath's reading, an important clue for holding that every instance of rational agency involves a claim to universality lies in Kant's view of self-conceit. Reath suggests that self-conceit entails a striving for universality that he takes to be an expression of the overall structure of practical rationality.²³ The main support for this reading is provided by an important passage where Kant describes self-conceit as self-love that "makes itself lawgiving and the unconditional practical principle" (KpV, AA 05: 74). In spite of the similarity with the formulations usually deployed with regard to the categorical imperative, that characterisation of self-conceit does entail that self-conceit involves a tacit claim to universality. Like in the heteronomous determination of the will the object "gives the law" (cf. GMS, AA 04: 434), self-conceit analogously gives a law by imposing the will of the subject on others.²⁴ Instead of an imperfect attempt at consistency in the moral realm, self-conceit is rather much more like a voluntarist demand put on others. Self-conceit, thus, need not be consistent, as the individual will wanting to direct everyone else can change, as it is not governed by a law in the first place. Thus this tendency can hardly entail a claim to universality that can be regarded as formally continuous with the universal validity of moral principles.

Furthermore, if the Thesis should hold for Kant, that would entail that the 'test' of universalisation is meant as an aspect of choice itself. In spite of the complexities

²³ Cf. Reath, *Did Kant Hold*, 251f.

²⁴ See on this, and for a full account of Kant's view of self-conceit, Moran, Kate: *Delusions of Virtue: Kant on Self-Conceit*. In: *Kantian Review* 19 (2014): 419-447, 436f.

regarding the role of universalisation procedures in Kant's view, to my present purposes it suffices to point out a significant difference separating Kant's conception from the implications of the Kantian version of the Thesis. The universalisation test is not presented as an aspect of rational volition as such, but only of moral volition specifically, that is, as the way how an empirically unconditioned principle can determine a content applying the concept of moral good, as becomes clear especially in the Typic. Kant's remarks there are not about how rational volition generally is structured, but about how a purely formal principle can guide volition. Through universalisation, the moral law functions like a compass (cf. GMS, AA 04: 404), orienting the choice towards moral actions.

The Thesis, thus, does not seem to play the role of a fundamental formal principle in Kant's general view of agency, since it does not hold for morally defective maxims, which are not universalisable and cannot be taken to manifest representations of goodness. For Kant, morally defective maxims do not arise from representations of the good, but from pleasure and interest. A morally defective maxim as a "mere maxim" is "a subjective principle on which we might act if we have the propensity and inclination" (GMS, AA 04: 425). The continuity between empirically conditioned and pure determination of the will lies rather in pleasure than in rational beliefs about goodness, as also moral maxims entail a kind of pleasure. In Kant's view, the Thesis only applies to the determination of the will *through the moral law*. Autonomy of the will as the capacity of a will to be a law to itself amounts to the capacity to determine itself only under the guise of the good. Here lies the grain of truth contained in that dogma of moral rationalism, for Kant. Wolff's descriptive principle thereby acquires in Kant's terms a different, normative status.

4

As has recently been observed with respect to his theoretical philosophy, Kant "not only flirts but also struggles with the rationalist legacy".²⁵ The same holds true for his moral thought as well, and his attitude to the Guise of the Good Thesis is a case in point. In this light Kant's (either apparent or only partial) commitment to the Thesis could be

²⁵ Pollok, Konstantin: *From the Clarity of Ideas to the Validity of Judgments: Kant's Farewell to Epistemic Perfectionism*. In: *Canadian Journal of Philosophy* 44 (2014), 18-35, 25.

explained as an attempt to accommodate the grain of truth of rationalism in a new philosophical framework. Now, if there is reason enough to ascribe the Thesis to Kant, this can happen, I suggest, only under several qualifications, that is, (1) If we make clear that no continuity in the objects of rational volition is entailed; (2) If we make clear that no prior evaluation of the object of choice is entailed, and that only the principles governing volition are relevant; (3) Finally and most importantly, if we make clear that the reference to something, i.e. to a possible action as 'good' holds only for universalisable maxims as opposed to morally defective maxims. If under these qualifications the Guise of the Good Thesis can still be attributed to Kant in a meaningful way, it is probably mainly an issue of terminology.²⁶

²⁶ I should like to thank Andy Reath and Jens Timmermann for discussions on these issues before and during the Kant Congress in Vienna.