

University of Groningen

## Between 'Indubitably Certain' and 'Quite Detrimental' to Philosophy

Carvalho, Vinicius

*Published in:*  
 Kantian Review

*DOI:*  
[10.1017/S1369415423000316](https://doi.org/10.1017/S1369415423000316)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
 Publisher's PDF, also known as Version of record

*Publication date:*  
 2023

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*  
 Carvalho, V. (2023). Between 'Indubitably Certain' and 'Quite Detrimental' to Philosophy: Kant on the Guise of the Good Thesis. *Kantian Review*, 28(4), 537-553. <https://doi.org/10.1017/S1369415423000316>

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*



ARTICLE

# Between ‘Indubitably Certain’ and ‘Quite Detrimental’ to Philosophy: Kant on the Guise of the Good Thesis

Vinicius Carvalho

University of Groningen, Groningen, The Netherlands  
Email: [v.pinto.de.carvalho@rug.nl](mailto:v.pinto.de.carvalho@rug.nl)

## Abstract

Kant clearly endorses some version of the ‘old formula of the schools’, according to which all volition is *sub ratione boni*. There has been a debate whether he holds this only for morally good actions. I argue that a closer look at the distinction between the good and the agreeable does not support this conclusion. Considering Kant’s account of the detrimental and the correct use of this thesis, I argue that rational beings always will *sub ratione boni*, even when they act immorally, because they act on principles. I argue that Kant’s accounts of self-love and self-conceit support this view.

**Keywords:** Immanuel Kant; guise of the good; self-love; self-conceit; Kantian ethics; Achenwall; *sub ratione boni*

## 1. Introduction

In the *Critique of Practical Reason*, Kant discusses what we now call the ‘guise of the good’ thesis (hereafter GG),<sup>1</sup> according to which whatever we pursue is conceived by us to be good. Calling it ‘an old formula of the schools’, he expresses it as the proposition ‘*nihil appetimus, nisi sub ratione boni; nihil aversamus, nisi sub ratione mali* [we desire nothing except under the form of the good; nothing is avoided except under the form of the bad]’ and explains that once it is properly understood, the thesis ‘is indubitably certain and at the same time quite clearly expressed’ (*KpV*, 5: 59–60).<sup>2</sup> Kant thus clearly endorses some version of GG. What is also certain is that there is a sharp divide in the literature when it comes to which version Kant endorsed. On the one hand, some contend that Kant held GG only in the case of *morally* good actions and that the thesis ‘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’ (Bacin 2018: 1712). I call this the *weak* GG reading of Kant, for it holds that *only* morally good actions are performed under the guise of the good. On the other hand, some argue that Kant holds GG for actions *tout court*. Sergio Tenenbaum, for

© The Author(s), 2023. Published by Cambridge University Press on behalf of Kantian Review. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

instance, argues that '[i]n acting badly, I still must represent the object of the faculty of desire as good, even if in a confused manner. On Kant's view, we must attribute that all our actions are pursued under the guise of the good as a condition of their imputability' (2021: 77).<sup>3</sup> I call this the *strong* GG reading of Kant, for it holds that all action is willed under the guise of the good.

This disagreement is fostered by Kant's twofold approach to GG. He thinks that while on a certain formulation the old formula is 'indubitably certain', it can also be 'very detrimental (*nachtheiligen*) to philosophy' by virtue of an ambiguity in the Latin concepts of *bonum* and *malum* (*KpV*, 5: 59). In this article, I first explain the sense in which GG can be detrimental to philosophy, using Gottfried Achenwall as an example of a philosopher who committed the mistake Kant claims to have avoided (section 2). As we will see, Kant argues that previous defenders of GG fell prey to an ambiguity in the Latin *bonum*: they failed to see that it can either mean 'good' (*das Gute*), that is, something of objective value, or 'well-being' (*das Wohl*), which denotes what satisfies by means of pleasure and is conducive to a state of agreeableness.<sup>4</sup> Having only the latter meaning at hand, their version of GG proclaimed that we can only will an action if we expect to gain pleasure from it. Yet this implied, for Kant, that 'the possibility of even thinking of a pure practical law was already removed in advance'; their supreme principle of morality 'was in every case heteronomy and they had to come unavoidably upon empirical conditions for a moral law' (*KpV*, 5: 63–4). I shall return to these claims in the next section.

In section 3, I argue that Kant holds GG to be indubitably certain because he understands it as the thesis that rational agents take their choices to be justified by rational principles. I explain that Kant takes 'good' to mean actions that are commanded by principles of reason, and since principles of reason command finite rational agents by means of imperatives, I use his account of hypothetical and categorical imperatives to show that he identifies two ways in which actions can be supported by principles, and thus commanded as *good* actions. I argue that in hypothetical imperatives, a principle of reason commands an action because it is deemed necessary as a means to an end. The action is good, though conditionally. In the case of categorical imperatives, however, the moral law justifies the action to a fuller extent, since the end being pursued is also subject to a test of rational scrutiny: everyone is justified (or has reason) to pursue it, since the maxim can be willed at the same time as a universal law. As we will see, this supports the strong GG view, for it establishes that *every* action, as long as the agent acts under the representation of a principle, is willed under the guise of the good.

Finally, I consider the objection that Kant's moral psychology suggests that agents often will an action not because it is good but simply because it is pleasurable (section 4). On the basis of textual evidence, I argue that his accounts of *self-love* and *self-conceit* show that even in cases of immoral action, agents conceive of their choices as supported by principles, and thus as good. These considerations further support the strong GG reading.

## 2. Achenwall and the detrimental use of GG

At least two reasons count in favour of using Achenwall to illustrate what Kant took to be a detrimental use of GG. First is the fact that Kant used Achenwall's *Natural Law* as

the textbook on which he based his own lectures on the subject, and the lecture transcript we have available (the Feyerabend Lectures on Natural Law) dates to the time Kant was writing the *Groundwork*. This gives us reason to consider Achenwall a major philosophical interlocutor in Kant's practical philosophy.<sup>5</sup> Second, Achenwall's formulation of the guise of the good thesis and his identification of the good with the pleasurable consequences of actions are remarkably clear and straightforward, making them suitable for a helpful comparison with Kant's view in the next section. As I hope to make clear, Kant's acceptance of GG includes a major critique of the concept of the good, such that he ultimately changes the order of determination between the will and the good. Whereas on the standard version of the ancient thesis the good determines the will, for Kant the will determines the good by means of its very own principles.

Achenwall begins his *Prolegomena to Natural Law* by dividing the soul into two faculties: the cognitive, through which the soul 'knows', and the appetitive, 'by which it strives for that which we represent to ourselves as good, and hence avoids what we represent as bad' (§2). After establishing that the higher faculty of appetite is the will (*voluntas*), Achenwall contends that 'every act of will (volition and nolition) stems from striving for some good [*boni*] or avoiding something bad [*mali*]' (§5, orig. emphasis). More precisely, he thinks that whenever one represents an object as an end to be achieved, one represents it as good, and whenever one represents an object as something that is to be avoided, one represents it as bad. Because it is possible to be mistaken when representing something as good or bad (either by representing a good as a bad or vice versa), Achenwall further introduces the concept of a *true good* (*veri boni*): something that is represented as good and *that is indeed good*. The point here is that Achenwall posits a necessary relation between evaluation and motivation: he claims that even though agents might err in their representation of objects as good or bad, whenever they represent something as a goal to be achieved – whenever they will something – it is because the object at least *appears* to be good. Achenwall thus clearly endorses a version of the guise of the good thesis.

His next step is to relate the concept of a true good to that of an obligation by means of the concept of a *law*. For Achenwall, a law is a proposition that expresses a true obligation for a free agent by positing a true good 'as the goal of the obligated subject' (§§14, 15). In other words, every law commands the agent to pursue something that is represented as a true good.<sup>6</sup> The important thing to ask now is: what is the meaning of *bonum* for Achenwall, and what things are good? He addresses this issue in §20 (original emphasis):

[T]he means of creating an obligation and law in general consists in linking someone's action that is to be executed freely to some good or bad consequence and proposing, i.e. representing, that consequence to the person to be obligated. For in this way it will become necessary for him to direct his free action in accordance with the proposed consequence – either a good one for which he hopes or a bad one which he fears. To that extent it can be admitted that every *means of obligating* consists in *proposed hope or fear*.

As this passage indicates, good and bad always refer to good or bad *consequences*. What is more, Achenwall explains that laws are 'armed' with divine rewards and punishments, the idea being that what makes a consequence good is its being

rewarded by God, and what makes a consequence bad is its being punished (§55). Stressing the connection with GG, this allows us to conclude that for Achenwall the incentive behind all intentional action is either the obtaining of good consequences in the form of pleasant divine rewards (because this is what 'good' means) or avoiding bad consequences in the form of divine punishments that prompt disgust (for this is what 'bad' means).<sup>7</sup> 'Human nature', according to Achenwall, 'is such that it can only be obligated through good or bad consequences' (§55).

Why does Kant think that this is a detrimental use of GG? The problem is not with the claim that all volition aims at the good and avoids the bad, but rather with the meaning of 'good' and 'bad'. Since for Achenwall 'good' ultimately means 'pleasant divine reward', his acceptance of GG implies that any act of volition is ultimately motivated by seeking pleasure or avoiding disgust. This is a problem, according to Kant, because it excludes the possibility of a priori practical laws (*KpV*, 5: 64), and it is easy to see why. If all volition aims at the good, and if the good is conceived as a sort of pleasure, then all volition is ultimately determined by the idea of a pleasing object. Yet an a priori practical law is precisely a law that determines the will *immediately*, i.e. without any object as the ground of determination (*KpV*, 5: 27). This version of GG is detrimental to philosophy because to hold that all volition aims at the good, and to think that what is good is what is pleasurable, is to deny the very possibility of practical reason's being pure, since in this picture there must be a pleasurable object that determines the will if there is to be any volition at all.

What is more, Kant maintains that all other philosophers who held GG defined the good in terms of the pleasurable. Even though they disagreed about what constitutes pleasure (e.g. happiness, perfection, etc.), the moral principle they found was inevitably heteronomous: it always had to command conditionally, in terms of duty's being a means to pleasure, since the only object of practical reason is pleasure.<sup>8</sup>

Yet how exactly did Achenwall and many others commit this mistake? Kant's explanation is that they fell prey to the ambiguity of the Latin concepts of *bonum* and *malum*. He remarks:

[T]he German language has the good fortune to possess expressions which do not allow this difference to be overlooked. For that which the Latins denominate with a single word, *bonum*, it has two very different concepts and equally different expressions as well: for *bonum* it has *das Gute* [good] and *das Wohl* [well-being], for *malum* it has *das Böse* [evil] and *das Übel* [ill-being] . . . (*KpV*, 5: 59–60)

I will explain the difference between *das Gute* and *das Wohl* in more detail in the next section. For now, we just need to see that, since there are two meanings of *bonum* and *malum*, Kant contends also that 'there are two very different appraisals of an action depending upon whether we take into consideration the *good* and *evil* of it or our *well-being* and *woe* (ill-being)' (*KpV*, 5: 59–60). In other words, to say that an action is good is to say something very different than that it is conducive to agreeableness or is pleasing. Previous defenders of GG nonetheless conflated these two meanings.

Furthermore, Kant points out that even the expression '*sub ratione boni*' itself can be understood in two different ways, depending on how one defines *bonum*. If the good is conceived in terms of the agreeable (as in Achenwall and the tradition before Kant), the expression is better rendered as 'we will something *in consequence of this idea* [i.e.

the idea of the good]', since here an object must first be represented as good (i.e. desirable because it brings pleasure), and this representation 'precede[s] volition as its determining ground' (*KpV*, 5: 59). We can say, in this case, that the will 'tracks' something that is recognized as of value because it promises agreeableness.<sup>9</sup>

But as we will see in the next section, the distinction between the good and the agreeable shows that actions can be good even though they cause pain, and that they can be evil although conducive to agreeableness. What is then the meaning of the expression? Instead of pursuing something *because* it has been represented as good, we actually 'represent to ourselves something as good when and *because we desire* (will) it' (*KpV*, 5: 59), thus rendering *sub ratione boni* as willing under the *idea* of the good, and not in consequence of it. Fundamentally, for Kant the will is not *after* objects of value: it *confers* value on objects and actions when it deems them necessary according to its own principles.

### 3. Kant and the indubitably certain use of GG

#### 3.1 The distinction between the good and the agreeable

To introduce the concept of an imperative in the *Groundwork*, Kant first distinguishes the good from the agreeable in the following passage:

Practically good (*gut*) ... is what determines the will by means of representations of reason, hence not from subjective causes, but objectively, i.e. from grounds that are valid for every rational being, as such. It is distinguished from the agreeable (*Angenehmen*), as that which influences the will only by means of sensation from merely subjective causes, which hold only for the senses of this and that one, and not as a principle of reason, which holds for everyone. (*GMS*, 4: 413, original emphasis)

The good is distinguished from the agreeable by the fact that what is good is objective, in the sense that a claim to an object's goodness is 'valid for every rational being as such', because it is decided by a principle of reason. On the other hand, what is agreeable is determined by sensation, such that the claim that an object is agreeable holds only for 'this or that one', that is, for the subject who asserts it and possibly (but not necessarily) for others who *happen* to have a similar sensible constitution such that the same object also pleases them.

Furthermore, Kant makes the same distinction in the second *Critique* by claiming that the use of language:

distinguishes the agreeable from the good and the disagreeable from the evil (*das Angenehme vom Guten, das Unangenehme vom Bösen unterscheidet*) and requires that good and evil always be appraised by reason and hence through concepts, which can be universally communicated, not through mere feeling, which is restricted to individual subjects and their receptivity. (*KpV*, 5: 58)<sup>10</sup>

In other words, claims to the effect that an object is agreeable are fundamentally subjective: they merely report that a certain object influences the faculty of desire by means of pleasure or pain.<sup>11</sup> On the other hand, since a *principle* of reason determines

what is good, good objects are to be recognized as such by all rational beings, and this is why '[w]hat we are to call good must be an object of the faculty of desire in the judgment of every reasonable (*vernünftigen*) human being, and evil an object of aversion in the eyes of everyone; hence for this appraisal reason is needed, in addition to sense' (*KpV*, 5: 61; see also *KU*, 5: 207–9). This shows that the ambiguity to which Kant draws attention is *not* between a moral and a nonmoral sense of good<sup>12</sup> – since *das Gute* applies to moral *and* nonmoral goodness – but rather between an objective and a subjective sense of good. It turns out that the latter is not really goodness, but rather agreeableness.

### 3.2 The good on Kant's account of imperatives

Analysing Kant's account of imperatives puts us in a favourable position to understand how principles of reason determine what is good. Imperatives express commands of reason, and they are issued by the recognition that a certain course of action is necessary according to a rational principle. Since rational principles determine what is good, every imperative depicts an action as *good*. This is clear in the following passage:

Because every practical law represents a possible action as good and hence, for a subject practically determinable by reason, as necessary, all imperatives are formulae for the determination of a will that is good in some way. Now, if the action would be good merely as a means to *something else*, the imperative is *hypothetical*; if the action is represented as good *in itself*, hence as necessary in a will that in itself conforms to reason, as its principle, then it is categorical. (*GMS*, 4: 414, original emphasis)

Imperatives determine an action for a will that is good *in some way*. In what ways can a will be good, then? Kant's discussion of imperatives in the *Groundwork* suggests that each kind of imperative expresses a different kind of goodness.<sup>13</sup> In the above passage Kant distinguishes hypothetical from categorical imperatives precisely by the way in which action performed under their command aims at realizing the good. In the former case, the agent wills an end and reason commands a certain action by recognizing it as practically necessary, on account of its being a necessary means to the achievement of that end. It is deemed good because it is recognized to be necessary according to what I will call here the 'instrumental principle' (IP): '[w]hoever wills the end also wills (in so far as reason has decisive influence on his actions) the indispensably necessary means to it that is in his control' (*GMS*, 4: 417).<sup>14</sup> As we have seen with the distinction between what is good and what is agreeable, to say that something is good is to make an objective claim. When it comes to good actions, this means two things: first, that the reasoning that leads an agent to judge that a certain action is good will lead other similarly situated agents to arrive at the same conclusion; second, that agents who do not face similar circumstances (i.e. they do not will the same end), thus agents for which those considerations will not lead to the conclusion that the action is good, will be in a position to at least comprehend why that action is the rational thing to do for those in the relevant



circumstances – they can see the goodness of it, in terms of the rationality of the action, even though they do not will the same end.

Consider Kant's example of a hypothetical imperative: if one wants to precisely divide a line into two equal parts, one ought to draw two intersecting arcs from the extremities of the line. This means that *anyone* who adopts that end ought to draw the intersecting arcs, since there is no other way to achieve the willed end; moreover, even those who could not care less about dividing a line into two equal parts are in a position to acknowledge that *if* they wanted to, they would have to draw two intersecting arcs from the extremities. Thus, 'drawing two intersecting arcs from the extremities of the line' is a *good* action when considered as a means to the end of 'dividing a line into two equal lines'. Consider another example: one wants to commit murder without being caught and discovers a certain poison that leaves no traces (say, ricin). Assume for the sake of argument that one knows that this is the only kind of poison that is suitable for one's purposes, such that to employ it is the only way to commit the murder without being caught. Because 'the use of ricin' is practically necessary for the end that is willed, it is deemed to be a *good* action, understood as the means to that end. As such, anyone who had the same end as our agent (i.e. killing without a trace) would be rationally constrained to use ricin, and even those who do not have this end can at least comprehend the rationale for using ricin given the adoption of that end, since it just is the indispensable means to it. This, then, is one way in which an action that is *good* makes a claim to the assent of every rational being.

Now, though it might seem surprising to use the concept of *goodness* to describe the selection of a poison to be used for murder,<sup>15</sup> it is important to keep in mind that a hypothetical imperative commands an action as good as a means to some possible or actual end. As such, Kant says, '[t]here is no question here whether the end is rational and good, but only what one must do in order to attain it' (GMS, 4: 415). In other words, when the choice of means is such that it best suits the promotion of the end, the imperative does command a good action, even if only relatively or conditionally. This should come as no surprise, for the selection of the best means to an end is governed by a rational principle (IP).

Of course, this does not mean that the goodness of the end should not be assessed. As long as the agent is rational, she will assess whether the pursuit of her end is also justified and good. Thus, while a hypothetical imperative commands an action that is conditionally good, a categorical imperative commands an action that is *absolutely* or *in itself* good, for the end one pursues is fully justified by a principle of reason. In other words, an action that is commanded by a categorical imperative is justified to a further level, because the principle on which the agent acts, her maxim, can be willed at the same time as a universal law (GMS, 4: 424). As such, *every* rational being, despite their private ends, has decisive reason to perform the action (or, in the case of actions that are prohibited, to omit it). And since goodness, as opposed to agreeableness, concerns that which holds as valid for every rational being, categorical imperatives can properly be said to command *absolute* or *unconditional* goodness.<sup>16</sup>

It follows that Kant's distinction between the good and the agreeable does not imply that actions are pursued under the guise of the good only when the will is determined by the moral law, as a weak GG reading would have it. To see this, consider Kant's examples that illustrate the distinction between *das Gute* and *das Wohl*: he wants to show that the distinction can explain why we sometimes say that an



action is good even though it involves pain (thus conducive to a state of *disagreeableness*). The first example concerns a man who undergoes surgery: because it involves pain it obviously feels like an ill, ‘but *through reason* he and everyone else pronounce it as good’ (KpV, 5: 61, my emphasis). In the second, he says that when someone gets a ‘sound thrashing’ (*einer tüchtigen Tracht Schläge*) for provoking and vexing others, the thrashing ‘is certainly an ill, yet *everyone would approve of it and take it as good in itself* even if nothing further resulted from it’ (KpV, 5: 61, my emphasis). These passages are difficult to interpret, but Kant’s wording suggests<sup>17</sup> that in the first example the action is mediately good – to undergo surgery is good for something, namely one’s health – whereas in the second the action is immediately good. In essence, the examples establish that these actions are judged to be good not because they bring about pleasure (the situations are designed precisely to show that they do not), but rather because the agent conceives of them as supported or justified by a principle of reason.<sup>18</sup>

Consequently, I take these considerations to count against the weak GG interpretation. Stefano Bacin, for example, claims that in ‘Kant’s view, the Thesis [GG] only applies to the determination of the will *through the moral law*’ (2018: 1712); that is, GG applies only when agents act from duty. Similarly, Louden (2021: 7) says that ‘[a]ny guise of the good that rational agents act on is one that has first been determined by the moral law’. But we have seen with the surgery example that the action is clearly pursued under the guise of the good, even though it is not determined by the moral law; as I noted before, it is presented as *mediately* good. According to the interpretation I propose here, that action is pursued under the guise of the good not because it is done from duty (as the weak GG reading would have it), but because the agent recognizes that it is made necessary by a principle of reason – in this case, the instrumental principle. I take this to be the essence of Kant’s claim that ‘we will nothing under the direction of reason (*nach Anweisung der Vernunft*) except insofar as we hold it to be good or evil’ (KpV, 5: 60). In other words, as long as the agent conceives of her choice as justified by a rational principle, the action is pursued under the guise of the good, regardless of whether it is determined by the moral law. As Kant remarks, the concepts of good and evil ‘stand under a practical rule of reason, which, *if it is pure reason, determines the will a priori with respect to its object*’ (KpV, 5: 67, my emphasis). This means that all principled action is willed under the guise of the good, and *if – and only if –* the moral law determines the will, the action is *absolutely and in itself* good (KpV, 5: 64).<sup>19</sup>

Nonetheless, Kant thinks that the general principle of self-love, as the principle of seeking one’s own happiness in all of one’s actions, is a powerful motivation for human beings (KpV, 5: 25). In fact, his account of moral motivation suggests that actions are motivated either by respect for the moral law or by inclinations serving self-love (GMS, 4: 400). Since a representation of pleasure is what ultimately motivates all these latter actions, do they not clearly have the *agreeable*, instead of the good, as their object? Is practical reason then not being merely ‘used’ to determine means in pursuit of pleasure? If this is the case, it appears that these actions are pursued under the guise of the agreeable, and not under the guise of the good, as the strong GG reading would have it. In the next section I further develop the contours of this objection and provide a reply to it.

#### 4. Nonmoral motivation and the guise of the good

Some passages seem to speak in favour of the objection outlined above. For instance, Kant argues that when we act on a hypothetical imperative:

[t]he end itself, the gratification that we seek, is in the latter case not a *good* but a *well-being*, not a concept of reason but an empirical concept of an object of feeling; but the use of means to it, that is, the action, is nevertheless called good (because rational reflection is required for it), not, however, good absolutely but only with reference to our sensibility, with respect to its feeling of pleasure and displeasure . . . (KpV, 5: 62, original emphasis)

This passage suggests that by acting on a hypothetical imperative one is most likely pursuing the satisfaction of some inclination.<sup>20</sup> There is a kind of goodness to the action because ‘rational reflection is required’ for the selection of the means to the end, yet we pursue the end *because* it promises pleasure – thus, it seems that we pursue the *means* under the guise of the good but the end itself under the guise of the agreeable. Consider a mundane example: Anna wants to eat the best pizza in town. She will have to act on a number of hypothetical imperatives to achieve this end: find out the best pizza place, figure out the best way to get there, which pizza will satisfy her craving, etc. These decisions will be called good insofar as they are rational decisions, following from the application of the instrumental principle. Nonetheless, they all ‘serve’ the relevant inclination, that of eating a nice pizza. Because what ultimately matters is the satisfaction of the inclination, Anna acts under the guise of the agreeable, not the good. As Kant says in the *Groundwork*, in such cases ‘reason states only the practical rule as to how to remedy the need (*Bedürfnis*) of inclination’ (GMS, 4: 413n.).

In the end, it appears that whenever we are not motivated by the feeling of respect for the law, we are motivated by inclination. And whenever we are motivated by inclination, we pursue the action because it promises pleasure. The action is in a sense good because reason selects the means to the end, but it is ultimately performed under the guise of the agreeable. I take passages like the above to provide the best support for the weak GG reading, on which, for Kant, only *morally* good actions are willed under the guise of the good.

My reply to this objection starts with the following consideration. Kant defines a rational being as having the ‘capacity to act *according to the representation* of laws, i.e. according to principles’ (GMS, 4: 412, original emphasis). Because this is a *capacity* (*Vermögen*), rational beings can fail to act on principles. The point, however, is that only when an action is derived from the representation of principles can we say that it is a product of practical reason, or the will.

In the same vein, consider the following passage, where Kant discusses how incentives influence choice:

Freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way

can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice (of freedom). (R, 6: 23–4, original emphasis)

It is not my intention here to discuss the intricacies of the Incorporation Thesis (see Allison 1990; Schapiro 2011). I merely want to point out that, on a standard reading of this passage, inclinations provide only incentives (*Triebfedern*) and cannot be acted upon unless the agent *chooses* to do so. One chooses to act on an incentive by taking it as a principle ('making it a universal rule', as Kant says). Presumably, even when the agent chooses an action just because it brings pleasure, she does so on the basis of a principle that is taken to justify that choice ('I will always choose what brings me pleasure', for instance).<sup>21</sup> After all, as we have just seen, to be a proper product of the will, an action must have been derived from the representation of a principle. What is more, it seems that as a *general* principle, self-love is the principle that structures all instances of nonmoral choice (see Reath 2006). I will develop this thought below.

We concluded in the last section that when the agent considers her choice to be justified, or supported by a principle, she does act under the guise of the good, as principles of reason command actions as 'practically necessary, i.e. as good' (GMS, 4: 412). With this in mind, in what follows I argue that Kant's accounts of self-love and, especially, of self-conceit provide the textual evidence we need for the claim that even when agents pursue agreeable ends, they still conceive of their actions as justified by a principle. This shows that they will under the guise of the good.

#### 4.1 Self-love under the guise of the good

In the second *Critique*, Kant defines self-love and self-conceit in the following passage:

All the inclinations together (which can be brought into a tolerable system and the satisfaction of which is then called one's own happiness) constitute regard for oneself (*solipsismus*). This is either the self-regard of *love for oneself*, a predominant *benevolence* toward oneself (*Philautia*), or that of *satisfaction with oneself* (*Arrogantia*). The former is called, in particular, *self-love*; the latter, *self-conceit*. Pure practical reason merely *infringes upon* self-love, inasmuch as it only restricts it, as natural and active in us even prior to the moral law, to the condition of agreement with this law, and then it is called *rational self-love*. (KpV, 5: 73)

This passage suggests that all actions with inclinations as their incentive come down to a form of either self-love or self-conceit. The former is characterized as a 'propensity to make oneself as having subjective determining grounds of choice into the objective determining ground of the will in general' (KpV, 5: 74). In other words, self-love consists in taking the promise of pleasure and happiness from acting on an inclination as a sufficient reason for choice. The fact that this propensity must take the form of an 'objective determining ground of the will' indicates that there must be some kind of principle on which the agent relies to make her choice; after all, deriving actions from principles is the *modus operandi* of the will (GMS, 4: 412). It is not surprising, then, that in discussing how nonmoral choice works, Kant says that when

one acts on inclinations one follows the ‘general principle of self-love’ (*allgemeine Princip der Selbstliebe*) (*KpV*, 5: 22). I take this to mean that in all instances of nonmoral action, that is, whenever inclinations are taken up as incentives, it is the general principle of self-love that ultimately guides the agent’s choice by highlighting certain courses of action as worthy of pursuit.

More precisely, the thought here is that, although it may appear at first that inclinations dictate how agents act, Kant’s portrayal of self-love as a principle shows that on his account one must still *choose* to pursue the ends urged by inclinations. And it is this act of choice or incorporation that makes the action be pursued under the guise of the good. In fact, what makes ends suggested by inclinations worthy of pursuit is not simply the fact that they satisfy the relevant inclination. Arguably, there could be different ends, the achievement of which would satisfy a given inclination, and further, finite rational beings are constantly subjected to a myriad of inclinations (*GMS*, 4: 417–19). As such, the ends suggested by inclinations need to be structured, so to say. I take Kant’s view to be that the principle of one’s own happiness is what ranks all these potential ends for action by assessing how much an end contributes to the agent’s opaque conception of happiness, ‘how much and how great satisfaction’ (*KpV*, 5: 23) it is expected to deliver. It is this assessment that ultimately provides the agent with a reason to pursue any end suggested by inclination. Thus, inclinations suggest the pursuit of certain ends, but one decides or chooses to pursue them only after the assessment of the principle of one’s own happiness. This principle, as any other practical principle in fact, marks actions as worthy of pursuit, or as good. As such, we could say that even when one acts on an inclination one acts under the guise of the good. But to be more precise, it is never the case that one acts *on an inclination*. One acts on the *principle* of self-love. We must not lose sight of the fact that Kant defines the will not as the capacity to follow incentives (inclinations or the feeling of respect for the law), but to act on the representation of *principles*, or laws. This account of the structure of nonmoral choice thus shows that in satisfying an inclination one acts on a principle which rules the action as a justified and good thing to do to further one’s happiness. And because the action is conceived as supported by a principle, it means that it was pursued under the guise of the good, for only good and evil can be appraised by reason (*KpV*, 5: 58).<sup>22</sup>

Self-love does not necessarily conflict with the moral law, however. Since the drive to satisfy inclinations is natural to finite rational beings, the moral law cannot annihilate self-love, but it can *restrict* it. Under what Kant calls rational self-love (*vernünftige Selbstliebe*), one acts on inclinations only *on the condition* that one’s maxim does not conflict with the moral law. This counts as a case where the action is willed under the guise of the good for the same reason that applies to the previous case: it is still based on the general principle of self-love. Now, although Tenenbaum and I both argue for the strong GG reading, I think it is important to point out a problem in his argument for the claim that actions under rational self-love are willed under the guise of the good. In defending his reading, Tenenbaum points to a passage where Kant says that when one limits one’s ‘maxim based on inclination in order to afford it the universality of a law’, one makes the maxim ‘suitable for pure practical reason’ (*KpV*, 5: 35). He takes this to be an appropriate description of what happens when the moral law succeeds in restricting self-love, that is, when a maxim that is originally supported by an inclination becomes suitable or appropriate (Kant’s term is

*angemessen*) for pure practical reason. And because ‘the only object of pure practical reason is the good’ (Tenenbaum 2021: 82), Tenenbaum concludes that action under rational self-love is proved to be willed under the guise of the good. I disagree with this strategy because I think Tenenbaum fails to notice that the only object of pure practical reason is not the good *simpliciter*, but the moral or absolute good. Accordingly, Tenenbaum’s reading leaves us with the conclusion that rational self-love has as its object the unconditional or moral good. Yet this conclusion is unsustainable, for it runs counter to Kant’s distinction between acting *in conformity with* duty and acting *from* duty (*GMS*, 4: 397), or between following the *letter* of the law and following its *spirit* (*KpV*, 5: 72). Furthermore, the passage used by Tenenbaum does not figure in Kant’s explanation of rational self-love, but rather is part of his argument for the duty of promoting the happiness of others, and this is what warrants Kant’s talk of the object of pure practical reason in that context. In sum, we should conclude from this discussion that only actions with the incentive of respect for the moral law as their determining ground are objects of pure practical reason and thus are willed under the guise of the absolute (moral) good.<sup>23</sup> Actions performed under rational self-love are willed under the guise of the good simply because they follow from the application of the principle of self-love. And whenever actions follow from the application of a principle, they are willed under the guise of the good.

#### 4.2 Self-conceit under the guise of the good

The structure of self-conceit reveals more clearly that agents still act under the guise of the good when they act immorally. Whereas in self-love the agent tries to justify acting on an inclination on a particular occasion, in self-conceit this propensity is taken to a further level. Self-conceit does not concern occasional transgressions, but instead amounts to a complete inversion of the principle that is regarded as objectively valid.<sup>24</sup> One who falls prey to self-conceit treats the satisfaction of one’s happiness as the objective principle to be followed – holding that acting on inclination is a principle that is valid not merely for oneself but for everyone, not only on a couple of occasions but *systematically*: self-conceit is self-love’s making ‘itself lawgiving and the unconditional practical principle’ (*KpV*, 5: 74).

Now, if self-conceit makes the principle of self-love the unconditional practical principle, the agent must somehow conceive of actions derived from the principle of self-love as unconditional *objects* of practical reason. Yet the unconditional object of practical reason is the ultimate kind of good, the moral good, in the form of the ‘fitness of maxims for giving universal law’ (*KpV*, 5: 74). This means, therefore, that self-conceit is a misconceived representation of happiness as the unconditional object of practical reason – happiness as *absolutely* good. Very tellingly, this is the exact dialectic that we find in the crucial paragraph where Kant describes self-love and self-conceit. In explaining these propensities, Kant quite pertinently begins the paragraph by talking about the ‘preceding chapter’, that is, the chapter on good and evil as objects of practical reason. He recalls that what makes an object absolutely good is the form of the maxim by which it is pursued, that is, whether the maxim is fit for holding as a universal law. ‘However’, he continues:

[W]e find our nature as sensible beings so constituted that the matter of the faculty of desire (objects of inclination, whether of hope or fear) first forces itself upon us, and we find our pathologically determinable self, even though it is quite unfit to give universal law through its maxims, nevertheless striving antecedently to make its claims primarily and universally valid, just as if it constituted our entire self. This propensity to make oneself as having subjective determining grounds of choice into the objective determining ground of the will in general can be called *self-love*; and if self-love makes itself lawgiving and the unconditional practical principle, it can be called *self-conceit*. (KpV, 5: 74, original emphasis)

This entire paragraph is designed to show that, while only the moral law is the unconditional practical principle by which the object of the will is absolutely good, self-conceit is such that it treats self-love, that is, the propensity to take one's inclinations as sufficient reasons for action, as if it were the unconditional practical principle.<sup>25</sup> It legislates the principle of attending to one's own happiness, even though we know that this principle cannot hold as a law (see Theorems I–II in the KpV). I take this to explain, at least partly, why Kant calls self-conceit a *delusion* (*Wahn*)<sup>26</sup> (KpV, 5: 75): self-conceit tries to elevate maxims of self-love to the category of moral maxims, illusorily treating happiness as if it were absolutely good. This is the reason why the moral law, 'which alone is truly objective', must infringe on self-conceit without end, for the latter 'prescribes as laws the subjective conditions of self-love' (KpV, 5: 75, my emphasis). This shows that in self-conceit agents still act under the guise of the good. More precisely, they act under the guise of the *unconditional* good by completely misrepresenting what the unconditional good is.

Thus, in short, in the case of self-love agents use a principle to justify an exception to a rule they recognize as authoritative. Self-conceit is particularly worrisome, however, for it elevates the principle of self-love to the status of law. It is as if the agent engaged in self-conceit regards the strategy of attending only to 'how intense, how long, how easily acquired, and how often repeated this agreeableness is' (KpV, 5: 23) as if it were approved by *pure* practical reason. Because in both cases agents act *nach Anweisung der Vernunft* – that is, they act on principles – it can properly be said that they act under the guise of the good.

## 5. Concluding remarks

In this article, I have provided a new set of considerations in support of the view that for Kant all volition is *sub ratione boni*, which I dubbed the *strong GG thesis*. The first step is to understand what Kant took to be the detrimental use of GG. He thinks that while previous philosophers were right in holding that all volition is *sub ratione boni*, their mistake was to define *bonum* in terms of the agreeable. As such, their claim that all volition is under the guise of the good was in fact a claim to the effect that, whatever we intentionally do, we ultimately do it because we are interested in the pleasure that it will bring us. And this renders moral principles necessarily heteronomous: all actions are ultimately performed with a view to their pleasurable consequences. Kant is adamant that by placing 'the practical concepts of good and evil *merely* (*blos*) in experiential consequences (so-called happiness)', philosophers were

bound to an ‘empiricism of practical reason’ (*KpV*, 5: 70, my emphasis). This is why such a use of the old formula of the schools is so detrimental to philosophy.

By pointing out the ambiguity of the Latin concepts, Kant can give a meaning to GG that renders it indubitably certain: since the good is determined by reason, the old formula amounts to the claim that rational beings take their choices to be justified by a principle. When the action is a necessary means to a willed end, it is justified by the instrumental principle. And if the action is necessary even without further reference to the agent’s ends, it is justified by practical reason’s fundamental principle, which states the conditions under which an action is *absolutely* good. What is more, the structure of nonmoral action reveals that for Kant, in order to act at all, rational agents must conceive of their choices as supported by principles, since that is the nature of practical reason. In a delusional move, they might go so far as to treat the principle of self-love as if it were the unconditional principle of practical reason.

**Acknowledgements.** I am grateful to the Dutch Research Council (NWO) for funding this research. I thank the audience at the 2022 UK Kant Society Conference and the two referees from *Kantian Review* for their helpful comments. I am grateful to Janis Schaab, Nicole Martinazzo, Marijana Vujosevic, Leon van Rijsbergen, Mike Gregory, Suzanne Jacobi, Fiorella Tomassini, Monique Hulshof, Carolyn Benson and especially Tobias Rosefeldt and Pauline Kleingeld for their detailed comments on earlier versions of this article.

## Notes

1 For classic criticisms of GG, see Stocker (1979) and Velleman (2000). For positive accounts, see Orsi (2015) and Tenenbaum (2007).

2 References to Kant’s works cite the abbreviation, volume and the *Akademie* page number. The first *Critique* is referenced following the standard A for the 1781 edition and B for the 1787 edition. Translations are from the *Cambridge Edition of the Works of Immanuel Kant* (ed. Paul Guyer and Allen Wood), except the *Groundwork*, for which I rely on the revised version of Mary Gregor’s original translation (ed. Jens Timmermann, Cambridge: Cambridge University Press, 2014). *KrV* = *Critique of Pure Reason*; *KpV* = *Critique of Practical Reason*; *V-Mo/Collins* = Collins Lectures on Ethics; *V-Mo/Vigilantius* = Vigilantius Lectures on Ethics; *GMS* = *Groundwork for the Metaphysics of Morals*; *R* = *Religion within the Boundaries of Mere Reason*; *V/Met-Mron* = Mronovius Lectures on Metaphysics; *KU* = *Critique of the Power of Judgement*; *V-PP/Herder* = Herder Lectures on Ethics; *V-Mo/Mron II* = Mronovius Lectures on Ethics; *TP* = *Theory and Practice*; *V-NR/Feyerabend* = Feyerabend Lectures on Natural Law.

3 See also Timmermann (2022) and Loudon (2021) for the weak GG reading. Reath (2015) advocates the strong GG reading while admitting that the texts are not decisive.

4 From now on I will treat the concept of agreeableness (*das Angenehme*) as synonymous with that of well-being (*das Wohl*). See *KpV*, 5: 60.

5 See Pauline Kleingeld’s introduction to Achenwall’s (2020a) *Prolegomena to Natural Law* and Paul Guyer’s introduction to Achenwall’s (2020b) *Natural Law*.

6 Furthermore, he remarks that ‘the use of the imperative in the practical disciplines indicates that a man is being obligated’ (§23), from which we can conclude that imperatives in the practical disciplines always command an action that is represented as *good*. As we shall see in section 3, Kant also maintains this relation between imperatives and the good.

7 To illustrate, the consequences that prompt disgust include sadness, grief, terror and so on. See §§68–70 for Achenwall’s full description of pleasure and disgust.

8 It is also instructive to see how Baumgarten (2013) makes the same mistake. He argues: ‘What pleases me I intuit as good, under the form of the good, and what displeases me I intuit as evil, under the form of evil.’ (*‘Quod placet, intueor, ut bonum, sub ratione boni, §100, quod displicet, intueor, ut malum, sub ratione mali, §146, Metaphysica, §651.*) Even more tellingly, he argues: ‘What I desire ... is pleasing. ... I do not desire things ... that are utterly not pleasing’ (*‘Quae appeto ... placet. ... quae prorsus non placent, ... non appeto’, Metaphysica, §664.*) That is, Baumgarten clearly identifies the good with what pleases.



Furthermore, since for Baumgarten what pleases is the intuition of perfection (*Metaphysica*, §665), it is clear that he is one of the ‘otherwise acute’ men whom Kant criticizes in *Remark I* of the *Critique of Practical Reason* for distinguishing the lower and the higher faculty of desire in terms of the *origin* of the representation that is connected with the feeling of the pleasure, whether it comes from sensibility (e.g. having a meal) or from the understanding (e.g. attending a fine speech). Kant thinks that this way of drawing the distinction, apart from conceptually denying the possibility of practical reason’s being pure, is disproved even by the commonest human reason (*KpV*, 5: 24). Note: I have altered the translation of the passages from Baumgarten, translating *ratione* as ‘form’ rather than ‘aspect’ to correspond with the English translation of the *KpV*, but I will still use *guise* of the good in referring to GG to match the use in the contemporary literature. Needless to say, the translation of *ratione* is a complicated matter.

**9** What if the desire for the good were conceived not in terms of pursuing a pleasurable object, but rather in terms of desiring an object cognized as good by the understanding, such as a determinate *concept* of happiness? Would Kant still regard this as a detrimental use of GG? The answer is yes, for this version of GG would still rule out there being practical laws for the pure use of practical reason. First of all, Kant insists that the concept of happiness is simply too indeterminate for finite rational beings, even for ‘the most insightful’ of them, to be able to establish what objects would make them happy (*GMS*, 4: 418). This means that the principle of one’s own happiness can only provide counsel as to what will *probably* make one happy, but it cannot command with the necessity of a law that which will make one happy. Secondly, even if one argues that happiness is desired because it is first recognized as good, Kant points out that at the individual level, what makes one happy or unhappy still ‘comes down to the particular feeling of pleasure or displeasure’ (*KpV*, 5: 25), so that happiness cannot be desired if not by the promise of pleasure associated with it. Kant further generalizes this point by arguing that whenever the concept of the good is used to derive laws for the will, ‘the criterion of good and evil could be placed in nothing other than the agreement of the object with our feeling of pleasure and displeasure’ (*KpV*, 5: 63). I thank an anonymous referee for calling my attention to this point.

**10** Kant’s discussion of the good as the object of practical reason in the *Critique of Practical Reason* is an attempt to answer the objections raised by Pistorius (1786) in his review of the *Groundwork*. For Pistorius, to say that a good will is one that acts out of respect for some law presupposes that following such a law is good. However, this must first be established by knowing what is good, for which, in turn, a definition of ‘good’ is needed. What is more, Pistorius thinks that, even if following a certain principle or law is what constitutes a good will, we still need to appeal to material considerations: it must be the case that following the law or principle results in good ends or objects. This is why Kant first starts out by providing a definition of ‘good’ by means of the distinction between the good and the agreeable. With this in hand, he further argues that material considerations, based on sensibility, can only determine what is agreeable, not good.

**11** What is essential for *well-being* and *ill-being* is that each ‘always signifies only a reference to our state of *agreeableness* or *disagreeableness*, of gratification or pain, and if we desire or avoid an object on this account we do so only insofar as it is referred to our sensibility and to the feeling of pleasure and displeasure it causes’ (*KpV*, 5: 60).

**12** See Loudon (2021: 5).

**13** The idea that there are *kinds* of goodness is prominent in the *Lectures on Ethics*, across many periods of Kant’s work. For instance, in the Collins lectures we read: ‘All imperatives are mere formulae of practical necessitation, and express a necessity of our actions under the condition of goodness ... The imperatives express objective necessitation, and since they are of three kinds, there are also three kinds of goodness’ (V-Mo/Collins, 27: 255–6). See also V-PP/Herder, 27: 4; V-Mo/Mron II, 27: 607; TP, 8: 282.

**14** By calling it the ‘instrumental principle’, my aim is to avoid taking a stance on whether there is such a thing as *the* Hypothetical Imperative. My only assumption is that the proposition ‘Whoever wills the end also wills (in so far as reason has decisive influence on his actions) the indispensably necessary means to it that is in his control’ (*GMS*, 4: 417) expresses a principle of practical reason for finite rational beings. It does not matter for the argument of this article whether this principle has the status of an imperative itself (Hill 1973 thinks that it does; Schroeder 2005 does not) or whether it is a rule of practical inference (see Schwartz 2010).

**15** Kant gives a somewhat similar example when discussing hypothetical imperatives of skill in the *Groundwork* (*GMS*, 4: 415). See also Beck (1960: 131–2).

16 See V-Mo/Collins, 27: 255–6, where Kant is reported as saying that categorical imperatives command *bonitas absoluta*.

17 The sentence leading up to the examples suggests that Kant is about to give different kinds of examples. He says: ‘But we can call something an ill which everyone must yet at the same time pronounce good, *sometimes* mediately but *sometimes* even immediately’ (KpV, 5: 61, my emphasis).

18 I am here in agreement with Korsgaard (1983: 181), who argues that, on Kant’s theory of action, rational beings always conceive of their choices as justified, and thus as good. See also Kleingeld (2016: 37), who claims that for Kant ‘the good is what we have reason to do, whether conditionally (dependent on antecedent aims) or unconditionally (morally, absolutely)’. For a recent criticism of Korsgaard’s view, see Timmermann (2022: 45).

19 See Kleingeld (2016: 37), who emphasizes that the good is the object of practical reason and the *absolutely* good is the object of *pure* practical reason.

20 I say most likely because hypothetical imperatives can (and usually are) necessary in order to fulfil moral obligations; e.g. if I promise to help a friend who is moving house, I will use many hypothetical imperatives in figuring out how to help.

21 Moreover, Kant says that while the practically good is what *determines* the will (*den Willen bestimmt*), the agreeable only *influences* it (*auf den Willen Einfluß hat*). This is a different way of making the point that humans have an *arbitrium liberum*, one that is influenced but not determined by sensible impulses. In this picture, a sensible incentive must be treated as a *motive* (*Bewegursache*) in order to determine the will (see KrV, A802/B830). In the Mrongovius Lectures on Metaphysics, Kant says that an *arbitrium liberum* is determined by *motives*, and the only motives are the concepts of the good (by means of which we desire things) and of evil (by means of which we avoid things). This supports the strong GG reading since, in other words, it suggests that while incentives influence the will, they must be taken as *good or evil*, and thus as motives, in order to determine it (see V/Met-Mron, 29: 896).

22 Similarly, Broadie and Pybus (1982: 410) argue that when we act immorally we still seek ‘to present to ourselves our effective motive as something acceptable to the judgment of reason’. Furthermore, Tenenbaum (2007: 10–11) remarks that it is characteristic of the guise of the good thesis that, in explaining intentional action, it aims to show ‘how an imperfectly rational agent came to the conclusion that a certain action was worth pursuing, even if the action in fact wasn’t or the agent should have known that it wasn’t’.

23 Although we both argue for a strong GG reading, Tenenbaum and I disagree about the role we assign to the Incorporation Thesis, at least to what seems to be the canonical way of interpreting it (following Allison 1990). Not only does Tenenbaum’s argument for a strong GG reading make no use of it, but he in fact also thinks that ‘we cannot explain Kant’s commitment to the guise of the good in its fullest extent [that is, per a strong GG reading] this way’ (2021: 87). Tenenbaum argues that the epistemic priority of the moral law vis-à-vis the consciousness of freedom (the former as the *ratio cognoscendi* of the latter) cannot be explained by the standard reading of the Incorporation Thesis. Even if he is right, this does not affect my use of it in arguing for a strong GG thesis. I do acknowledge that we must not invert the *ratio cognoscendi* of our awareness of freedom: we know we are free not because we can choose whether or not to follow an inclination, but rather because we are conscious of moral obligation. Yet it is still the case that rational beings are not simply determined to follow their inclinations; rather, they are free to choose to satisfy them or not, and this act of choice is based on a principle. The way the interpretation proposed here advances beyond Tenenbaum’s position is by showing the centrality of the Incorporation Thesis in arguing for a strong GG reading.

24 The view of self-conceit I develop here is in agreement with Tenenbaum (2021).

25 Bacin (2018: 1711) argues that there is no parallel between the alleged claim to universality of self-conceit and the universality of moral principles because he adopts the view that self-conceit consists *exclusively* in the imposition of the subject’s will on others. It has recently been argued, however, that this ‘interpersonal’ aspect of self-conceit either depends on the subject’s mistakenly taking self-love as his or her own supreme practical principle (Moran 2014, which Bacin surprisingly references) or is simultaneous with it (Russell 2020).

26 See Moran (2014: 439) for the claim that self-conceit consists in an *illusion* of virtue. See also V-Mo/Collins, 27: 464, where Kant describes self-conceit as a ‘dream-like condition’.

## References

- Achenwall, G. (2020a) *Prolegomena to Natural Law*. Ed. P. Kleingeld, trans. C. L. Vermeulen. Groningen: University of Groningen Press.
- (2020b) *Natural Law: A Translation of the Textbook for Kant's Lectures on Legal and Political Philosophy*. Ed. P. Kleingeld, trans. C. L. Vermeulen, with an introduction by P. Guyer. London: Bloomsbury Academic.
- Allison, H. E. (1990) *Kant's Theory of Freedom*. New York: Cambridge University Press.
- Bacin, S. (2018) 'Under the Guise of the Good'. In V. L. Waibel, M. Ruffing and D. Wagner (eds), *Natur und Freiheit* (Berlin: De Gruyter), 1705–14.
- Baumgarten, A. G. (2013) *Metaphysics: A Critical Translation with Kant's Elucidations, Selected Notes and Related Materials*. Ed. and trans. C. D. Fugate and J. Hymers. New York: Bloomsbury.
- Beck, L. W. (1960) *A Commentary on Kant's Critique of Practical Reason*. Chicago: University of Chicago Press.
- Broadie, A. and Pybus, E. M. (1982) 'Kant and Weakness of Will'. *Kant-Studien*, 73, 406–12.
- Hill, T. E. (1973) 'The Hypothetical Imperative'. *Philosophical Review*, 82, 429–50.
- Kleingeld, P. (2016) 'Kant on "Good", the Good, and the Duty to Promote the Highest Good'. In T. Höwing (ed.), *The Highest Good in Kant's Philosophy* (Berlin: De Gruyter), 33–49.
- Korsgaard, C. M. (1983) 'Two Distinctions in Goodness'. *Philosophical Review*, 92, 169–95.
- Louden, R. B. (2021) 'Kant and the "Old Formula of the Schools"'. *Philosophical Explorations*, 24, 63–74.
- Moran, K. (2014) 'Delusions of Virtue: Kant on Self-Conceit'. *Kantian Review*, 19, 419–47.
- Orsi, F. (2015) 'The Guise of the Good'. *Philosophy Compass*, 10, 714–24.
- Pistorius, H. A. (1786) 'Rezension der Grundlegung zur Metaphysik der Sitten'. *Allgemeine Deutsche Bibliothek*, 66, 447–63.
- Reath, A. (2006) 'Hedonism, Heteronomy, and Kant's Principle of Happiness'. In *Agency and Autonomy in Kant's Moral Theory* (Oxford: Oxford University Press), 33–66.
- (2015) 'Did Kant hold that Rational Volition is Sub Ratione Boni?'. In M. Timmons, R. N. Johnson (eds), *Reason, Value, and Respect: Kantian Themes from the Philosophy of Thomas E. Hill, Jr.* (Oxford: Oxford University Press), 232–55.
- Russell, F. (2020) 'Kantian Self-Conceit and the Two Guises of Authority'. *Canadian Journal of Philosophy*, 50, 268–83.
- Schapiro, T. (2011) 'Foregrounding Desire: A Defense of Kant's Incorporation Thesis'. *Journal of Ethics*, 15, 147–67.
- Schroeder, M. (2005) 'The Hypothetical Imperative?'. *Australasian Journal of Philosophy*, 83, 357–72.
- Schwartz, J. (2010) 'Do Hypothetical Imperatives Require Categorical Imperatives?'. *European Journal of Philosophy*, 18, 84–107.
- Stocker, M. (1979) 'Desiring the Bad: An Essay in Moral Psychology'. *Journal of Philosophy*, 76, 738–53.
- Tenenbaum, S. (2007) *Appearances of the Good: An Essay on the Nature of Practical Reason*. New York: Cambridge University Press.
- (2021) 'Duality of Motivation and the Guise of the Good in Kant's Practical Philosophy'. *Philosophical Explorations*, 24, 75–92.
- Timmermann, J. (2022) *Kant's Will at the Crossroads: An Essay on the Failings of Practical Rationality*. Oxford: Oxford University Press.
- Velleman, J. D. (2000) 'The Guise of the Good'. In *The Possibility of Practical Reason* (Oxford: Oxford University Press), 99–122.

---

**Cite this article:** Carvalho, V. (2023). Between 'Indubitably Certain' and 'Quite Detrimental' to Philosophy: Kant on the Guise of the Good Thesis. *Kantian Review* 28, 537–553. <https://doi.org/10.1017/S1369415423000316>